R reference card
extracted from Tinn-R file Rcard.xml

A. Blejec

October 27, 2010

Contents

1 Introduction

A collection of R commands is available in Tinn-R. Here we will extract a list of functions from a XML file.

2 Reading XML file

XML tree is in the file xmlFile. First we will read the file:
Look at the node names.
We are interested in the second node
Convert xml tree into more R like object
Convert it to a `data.fame`
Exclude Programming examples

code description

3 Basic and help

?topic Documentation on topic

apropos('topic') The names of all objects in the search list matching the regular expression 'topic' the regular expression 'topic'

attach(x) Database x to the R search path; x can be a list, data frame, or R data file created with save. Use search() to show the search path

detach(x) x from the R search path; x can be a name or character string of an object previously attached or a package

dir() Show files in the current directory

example(command) Examples of command

help(package=mva) Help with (e.g.) package mva

help(topic) Documentation on topic

help.search('topic') Search the help system

help.start() Start the HTML version of help

library(x) Load add-on packages; library(help=x) lists datasets and functions in package x

ls() Show objects in the search path; specify pat='pat' to search on a pattern

ls.str() str() for each variable in the search path

methods(a) Shows S3 methods of a

methods(class=class(a)) Lists all the methods to handle objects of class a

options(...) Set or examine many global options; common ones: width, digits, error

str(a) Display the internal structure of an R object

summary(a) Gives a 'summary' of a, usually a statistical summary but it is generic meaning it has different operations for different classes of a

4 Data (creation)

array(x, dim=) Array with data x; specify dimensions like dim=c(3,4,2); elements of x recycle if x is not long enough

c(...) Generic function to combine arguments with the default forming a vector; with recursive=TRUE descends through lists combining all elements into one vector

cbind(...) Idem rbind(...) by columns

data.frame(...) Create a data frame of the named or unnamed arguments; data.frame(v=1:4, ch=c('a', 'b', 'c', 'd'), n=10); shorter vectors are recycled to the length of the longest

expand.grid() A data frame from all combinations of the supplied vectors or factors

factor(x, levels=) Encodes a vector x as a factor

from:to Generates a sequence; ':' has operator priority; 1:4 + 1 is '2, 3, 4, 5'

gl(n, k, length=n*k, labels=1:n) Generate levels (factors) by specifying the pattern of their levels; k is the number of levels, and n is the number of replications

list(...) Create a list of the named or unnamed arguments; list(a=c(1, 2), b='hi', c=3i);

matrix(x, nrow=, ncol=) Matrix; elements of x recycle

rbind(...) Combine arguments by rows for matrices, data frames, and others

rep(x, times) Replicate x times; use 'each=' to repeat 'each' element of x each times; rep(c(1, 2, 3), 2) is 1 2 3 1 2 3; rep(c(1, 2, 3), each=2) is 1 1 2 2 3 3

seq(along=x) Generates 1, 2, ..., length(x); useful for for loops

seq(from, to) Generates a sequence 'by=' specifies increment; 'length=' specifies desired length

5 Data (load, read, write and save)

data(x) Loads specified data sets

load() Load the datasets written with save

read.csv('file', header=TRUE) Idem read.table(file), but with defaults set for reading comma-delimited files

read.delim('file', header=TRUE) Idem read.table(file), but with defaults set for reading tab-delimited files

read.fwf('file', widths, header=FALSE, ...) Read a table of fixed width formatted data into a 'data.frame'; widths is an integer vector, giving the widths of the fixed-width fields

read.table('file') Reads a file in table format and creates a data frame from it; the default separator sep=' ' is any whitespace; use header=TRUE to read the first line as a header of column names; use as.is=TRUE to prevent character vectors from being converted to factors; use comment.char=' ' to prevent '#' from being interpreted as a comment; use skip=n to skip n lines before reading data; see the help for options on row naming, NA treatment, and others

read.table(file('clipboard'), header=T) Read a spreadsheet content from clipboard

save('file', ...) Saves the specified objects (...) in the XDR platformindependent binary format

save.image('file') Saves all objects

write(object, 'file') Writes object to fileName

write.table(dataFrame, 'file') Writes a table

write.table(x, file=' ', row.names=TRUE, ...) Prints x after converting to a data frame; if quote is TRUE, character or factor columns are surrounded by quotes (' '); sep is the field separator; eol is the end-of-line separator; na is the string for missing values; use col.names=NA to add a blank column header to get the column headers aligned correctly for spreadsheet input

6 Data (selection and manipulation)

choose(n, k) Computes the combinations of k events among n repetitions=n!/[(n-k)!k!]

cut(x, breaks) Divides x into intervals (factors); breaks is the number of cut intervals or a vector of cut points

match(x, y) Returns a vector of the same length than x with the elements of x which are in y (NA otherwise)

na.fail(x) Returns an error message if x contains at least one NA

na.omit(x) Suppresses the observations with missing data (NA) (suppresses the corresponding line if x is a matrix or a data frame)

prop.table(x, margin=) Table entries as fraction of marginal table

rev(x) Reverses the elements of x

sample(x, size) Resample randomly and without replacement size elements in the vector x, the option replace=TRUE allows to resample with replacement

sort(x) Sorts the elements of x in increasing order; to sort in decreasing order: rev(sort(x))

subset(x, ...) Returns a selection of x with respect to criteria (..., typically comparisons); if x is a data frame, the option select gives the variables to be kept or dropped using a minus sign

table(x) Returns a table with the numbers of the differents values of x (typically for integers or factors)

unique(x) If x is a vector or a data frame, returns a similar object but with the duplicate elements suppressed

which(x==a) Returns a vector of the indices of x if the comparison operation is true (TRUE), in this example the values of i for which x[i]==a (the argument of this function must be a variable of mode logical)

which.max(x) Returns the index of the greatest element of x

which.min(x) Returns the index of the smallest element of x

7 Dates and times

as.Date(s) Convert to the respective class; format(dt) converts to a string representation. The default string format is '2001-02-21'. These accept a second argument to specify a format for conversion. Some common formats are: %a, %A Abbreviated and full weekday name; %b, %B Abbreviated and full month name; %d Day of the month (01-31); %H Hours (00-23); %I Hours (01-12); %j Day of year (001-366); %m Month (01-12); %M Minute (00-59); %p AM/PM indicator; %S Second as decimal number (00-61); %U Week (00-53); the first Sunday as day 1 of week 1; %w Weekday (0-6, Sunday is 0); %W Week (00-53); the first Monday as day 1 of week 1; %y Year without century (00-99). Don't use; %Y Year with century; %z (output only.) Offset from Greenwich; -0800 is 8 hours west of; %Z (output only.) Time zone as a character string (empty if not available)

as.POSIXct(s) Convert to the respective class; format(dt) converts to a string representation. The default string format is '2001-02-21'. These accept a second argument to specify a format for conversion. Some common formats are: %a, %A Abbreviated and full weekday name; %b, %B Abbreviated and full month name; %d Day of the month (01-31); %H Hours (00-23); %I Hours (01-12); %j Day of year (001-366); %m Month (01-12); %M Minute (00-59); %p AM/PM indicator; %S Second as decimal number (00-61); %U Week (00-53); the first Sunday as day 1 of week 1; %w Weekday (0-6, Sunday is 0); %W Week (00-53); the first Monday as day 1 of week 1; %y Year without century (00-99). Don't use; %Y Year with century; %z (output only.) Offset from Greenwich; -0800 is 8 hours west of; %Z (output only.) Time zone as a character string (empty if not available)

8 Distributions

rbeta(n, shape1, shape2) Beta

rbinom(n, size, prob) Binomial

rcauchy(n, location=0, scale=1) Cauchy

rchisq(n, df) Pearson

rexp(n, rate=1) Exponential

rf(n, df1, df2) Fisher-Snedecor

rgamma(n, shape, scale=1) Gamma

rgeom(n, prob) Geometric

rhyper(nn, m, n, k) Hypergeometric

rlnorm(n, meanlog=0, sdlog=1) Lognormal

rlogis(n, location=0, scale=1) Logistic

rnbinom(n, size, prob) Negative binomial

rnorm(n, mean=0, sd=1) Gaussian (normal)

rpois(n, lambda) Poisson

rsignrank(nn, n) Wilcoxon's statistics

rt(n, df) Student (t)

runif(n, min=0, max=1) Uniform

rweibull(n, shape, scale=1) Weibull

rwilcox(nn, m, n) Wilcoxon's statistics

9 Graphical (low-level commands)

abline(a, b) Draws a line of slope b and intercept a

abline(h=y) Draws a horizontal line at ordinate y

abline(lm.obj) Draws the regression line given by lm.obj

abline(v=x) Draws a vertical line at abcissa x

arrows(x0, y0, x1, y1, angle=30, code=2) Idem segments(x0, y0, x1, y1), but with arrows at points (x0, y0) if code=2, at points (x1, y1) if code=1, or both if code=3; angle controls the angle from the shaft of the arrow to the edge of the arrow head

axis(side) Adds an axis at the bottom (side=1), on the left (2), at the top (3), or on the right (4); at=vect (optional) gives the abcissa (or ordinates) where tick-marks are drawn

box() Draw a box around the current plot

legend(x, y, legend) Adds the legend at the point (x,y) with the symbols given by legend

lines(x, y) Adds lines (the option 'type=' can be used)

locator(n, type='n', ...) Returns the coordinates (x, y) after the user has clicked n times on the plot with the mouse; also draws symbols (type='p') or lines (type='l') with respect to optional graphic parameters (...); by default nothing is drawn (type='n')

mtext(text, side=3, line=0, ...) Adds text given by text in the margin specified by side (see axis()); line specifies the line from the plotting area

points(x, y) Adds points (the option ′type=′ can be used)

polygon(x, y) Draws a polygon linking the points with coordinates given by x and y

rect(x1, y1, x2, y2) Draws a rectangle which left, right, bottom, and top limits are x1, x2, y1, and y2, respectively

rug(x) Draws the data x on the x-axis as small vertical lines

segments(x0, y0, x1, y1) Draws lines from points (x0, y0) to points (x1, y1)

text(x, y, labels, ...) Adds text given by labels at coordinates (x, y); a typical use is: plot(x, y, type=′n′); text(x, y, names)

title() Adds a title and optionally a sub-title

10 Graphical (parameters)

adj Controls text justification (0 left-justified, 0.5 centred, 1 right-justified)

bg Specifies the colour of the background (ex. : bg=′red′, bg=′blue′, . . . the list of the 657 available colours is displayed with colors())

bty Controls the type of box drawn around the plot, allowed values are: ′o′, ′l′, ′7′, ′c′, ′u′ ou ′]′ (the box looks like the corresponding character); if bty=′n′ the box is not drawn

cex A value controlling the size of texts and symbols with respect to the default; the following parameters have the same control for numbers on the axes, cex.axis, the axis labels, cex.lab, the title, cex.main, and the sub-title, cex.sub

col Controls the color of symbols and lines; use color names: ′red′, ′blue′ see colors() or as ′#RRGGBB′; see rgb(), hsv(), gray(), and rainbow(); as for cex there are: col.axis, col.lab, col.main, col.sub

font An integer which controls the style of text (1: normal, 2: italics, 3: bold, 4: bold italics); as for cex there are: font.axis, font.lab, font.main, font.sub

las An integer which controls the orientation of the axis labels (0: parallel to the axes, 1: horizontal, 2: perpendicular to the axes, 3: vertical)

lty Controls the type of lines, can be an integer or string (1: ′solid′, 2: ′dashed′, 3: ′dotted′, 4: ′dotdash′, 5: ′longdash′, 6: ′twodash′, or a string of up to eight characters (between ′0′ and ′9′) which specifies alternatively the length, in points or pixels, of the drawn elements and the blanks, for example lty=′44′ will have the same effect than lty=2

lwd A numeric which controls the width of lines, default 1

mar A vector of 4 numeric values which control the space between the axes and the border of the graph of the form c(bottom, left, top, right), the default values are c(5.1, 4.1, 4.1, 2.1)

mfcol A vector of the form c(nr, nc) which partitions the graphic window as a matrix of nr lines and nc columns, the plots are then drawn in columns

mfrow Idem mfcol, but the plots are drawn by row

pch Controls the type of symbol, either an integer between 1 and 25, or any single character within ' '

ps An integer which controls the size in points of texts and symbols

pty A character which specifies the type of the plotting region, 's': square, 'm': maximal

tck A value which specifies the length of tick-marks on the axes as a fraction of the smallest of the width or height of the plot; if tck=1 a grid is drawn

tcl A value which specifies the length of tick-marks on the axes as a fraction of the height of a line of text (by default tcl=-0.5)

xaxs Style of axis interval calculation; default 'r' for an extra space; 'i' for no extra space

xaxt If xaxt='n' the x-axis is set but not drawn (useful in conjunction with axis(side=1, ...))

yaxs Style of axis interval calculation; default 'r' for an extra space; 'i' for no extra space

yaxt If yaxt='n' the y-axis is set but not drawn (useful in conjonction with axis(side=2, ...))

11 Graphical (plotting)

add=FALSE Parameters are common to many plotting functions, if TRUE superposes the plot on the previous one (if it exists)

assocplot(x) Cohen-Friendly graph showing the deviations from independence of rows and columns in a two dimensional contingency table

axes=TRUE Parameters are common to many plotting functions, if FALSE does not draw the axes and the box

barplot(x) Histogram of the values of x; use horiz=FALSE for horizontal bars

boxplot(x) Box-and-whiskers plot

contour(x, y, z) Contour plot (data are interpolated to draw the curves), x and y must be vectors and z must be a matrix so that dim(z)=c(length(x), length(y)) (x and y may be omitted)

coplot(x y | z) Bivariate plot of x and y for each value or interval of values of z

dotchart(x) If x is a data frame, plots a Cleveland dot plot (stacked plots line-by-line and column-by-column)

filled.contour(x, y, z) Idem contour(x, y, z), but the areas between the contours are coloured, and a legend of the colours is drawn as well

fourfoldplot(x) Visualizes, with quarters of circles, the association between two dichotomous variables for different populations (x must be an array with dim=c(2, 2, k), or a matrix with dim=c(2, 2) if k=1)

hist(x) Histogram of the frequencies of x

image(x, y, z) Idem contour(x, y, z), but with colours (actual data are plotted)

interaction.plot(f1, f2, y) If f1 and f2 are factors, plots the means of y (on the y-axis) with respect to the values of f1 (on the x-axis) and of f2 (different curves); the option fun allows to choose the summary statistic of y (by default fun=mean)

main= Parameters are common to many plotting functions, main title, must be a variable of mode character

matplot(x, y) Bivariate plot of the first column of x vs. the first one of y, the second one of x vs. the second one of y, etc.

mosaicplot(x) Mosaic graph of the residuals from a log-linear regression of a contingency table

pairs(x) If x is a matrix or a data frame, draws all possible bivariate plots between the columns of x

persp(x, y, z) Idem contour(x, y, z), but in perspective (actual data are plotted)

pie(x) Circular pie-chart

plot(x) Plot of the values of x (on the y-axis) ordered on the x-axis

plot(x, y) Bivariate plot of x (on the x-axis) and y (on the y-axis)

plot.ts(x) If x is an object of class 'ts', plot of x with respect to time, x may be multivariate but the series must have the same frequency and dates

qqnorm(x) Quantiles of x with respect to the values expected under a normal law

qqplot(x, y) Quantiles of y with respect to the quantiles of x

stars(x) If x is a matrix or a data frame, draws a graph with segments or a star where each row of x is represented by a star and the columns are the lengths of the segments

stem(x) produces a stem-and-leaf plot of the values in 'x'

stripplot(x) Plot of the values of x on a line (an alternative to boxplot() for small sample sizes)

sub= Parameters are common to many plotting functions, sub-title (written in a smaller font)

sunflowerplot(x, y) Idem than plot() but the points with similar coorcoordinates are drawn as flowers which petal number represents the number of points

symbols(x, y, ...) Draws, at the coordinates given by x and y, symbols (circles, squares, rectangles, stars, thermometres or 'boxplots') which sizes, colours . . . are specified by supplementary arguments

termplot(mod.obj) Plot of the (partial) effects of a regression model (mod.obj)

ts.plot(x) Idem plot.ts(x) but if x is multivariate the series may have different dates and must have the same frequency

type='p' Parameters are common to many plotting functions, specifies the type of plot, 'p': points, 'l': lines, 'b': points connected by lines, 'o': id. but the lines are over the points, 'h': vertical lines, 's': steps, the data are represented by the top of the vertical lines, 'S': id. but the data are represented by the bottom of the vertical lines

xlab= Parameters are common to many plotting functions, annotates the axes, must be variables of mode character

xlim= Parameters are common to many plotting functions, specifies the lower and upper limits of the axes, for example with xlim=c(1, 10) or xlim=range(x)

ylab= Parameters are common to many plotting functions, annotates the axes, must be variables of mode character

ylim= Parameters are common to many plotting functions, specifies the lower and upper limits of the axes, for example with ylim=c(1, 10) or ylim=range(x)

12 Graphics (devices)

bitmap See ?Devices

dev.off() Shuts down the specified (default is the current) graphics device; see also dev.cur, dev.set see also dev.cur, dev.set

jpeg See ?Devices

pdf See ?Devices

pictex See ?Devices

png See ?Devices

postscript(file) Starts the graphics device driver for producing PostScript graphics; use horizontal=FALSE, onefile=FALSE, paper='special' for EPS files; 'family=' specifies the font (AvantGarde, Bookman, Courier, Helvetica, Helvetica-Narrow, NewCenturySchoolbook, Palatino, Times, or ComputerModern); 'width=' and 'height=' specifies the size of the region in inches (for paper='special', these specify the paper size)

ps.options() Set and view (if called without arguments) default values for the arguments to postscript

windows() Open a graphics window

x11() Open a graphics window

xfig See ?Devices

13 Graphics (lattice)

barchart(y x) Histogram of the values of y with respect to those of x

bwplot(y x) Box-and-whiskers plot

cloud(z x*y|g1*g2) 3d scatter plot

densityplot(x) Density functions plot

dotplot(y x) Cleveland dot plot (stacked plots line-by-line and columnby-column)

histogram(x) Histogram of the frequencies of x

levelplot(z x*y|g1*g2) Coloured plot of the values of z at the coordinates given by x and y (x, y and z are all of the same length)

parallel(x) Parallel coordinates plot

qq(y x) Quantiles to compare two distributions, x must be numeric, y may be numeric, character, or factor but must have two 'levels'

qqmath(x) Quantiles of x with respect to the values expected under a theoretical distribution

splom(x) Matrix of bivariate plots

stripplot(y x) Single dimension plot, x must be numeric, y may be a factor

wireframe(z x*y|g1*g2) 3d surface plot

xyplot(y x) Bivariate plots (with many functionalities)

14 Indexing (data frames)

x$name Column named 'name'

x[['name']] Column named 'name'

15 Indexing (lists)

x$name Element of the list named 'name'

x[['name']] Element of the list named 'name'

x[[n]] Nth element of the list

x[n] List with elements n

16 Indexing (matrices)

x[, c(1, 3)] Columns 1 and 3

x[, j] Column j

x['name',] Row named 'name'

x[i,] Row i

x[i, j] Element at row i, column j

17 Indexing (vectors)

x['name'] Element named 'name'

x[-(1:n)] Elements from n+1 to the end

x[1:n] First n elements

x[c(1, 4, 2)] Specific elements

x[-n] All but the nth element

x[n] Nth element

x[x %in% c('a', 'and', 'the')] Elements in the given set

18 Input and output

cat(..., file=' ', sep=' ') Prints the arguments after coercing to character; sep is the character separator between arguments

data.entry() Spreadsheet

download.file('url1') From internet

format(x, ...) Format an R object for pretty printing

print(a, ...) Prints its arguments; generic, meaning it can have different methods for different objects

read.table.url('url1') Remote input

scan(x) Read a vector x

sink(file) Output to file, until sink()

source('file') Run the commands in file

source(file('clipboard')) Run the commands in clipboard

url.show('url') Remote input

19 Math

acos

Arg(x) Angle in radians of the complex number

asin

atan

atan2

Conj(x) Complex conjugate

convolve(x, y) Compute the several kinds of convolutions of two sequences

cos

cov(x, y) Covariance between x and y, or between the columns of x and those of y if they are matrices or data frames

cummax(x) A vector which ith element is the maximum from x[1] to x[i]

cummin(x) A vector which ith element is the minimum from x[1] to x[i]

cumprod(x) A vector which ith element is the product from x[1] to x[i]

cumsum(x) A vector which ith element is the sum from x[1] to x[i]

diff(x) Lagged and iterated differences of vector x

exp

fft(x) Fast Fourier Transform of an array

filter(x, filter) Applies linear filtering to a univariate time series or to each series separately of a multivariate time series

Im(x) Imaginary part

intersect(x, y) 'set' function

is.element(el, set) 'set' function

log

log(x, base) Computes the logarithm of x with base base

log10

max(x) Maximum of the elements of x

min(x) Minimum of the elements of x

Mod(x) Modulus; abs(x) is the same

mvfft(x) FFT of each column of a matrix

pmax(x, y, ...) A vector which ith element is the maximum of x[i], y[i], . . .

pmin(x, y, ...) A vector which ith element is the minimum of x[i], y[i], . . .

prod(x) Product of the elements of x

range(x) Idem then c(min(x), max(x))

rank(x) Ranks of the elements of x

Re(x) Real part of a complex number

round(x, n) Rounds the elements of x to n decimals

scale(x) If x is a matrix, centers and scales the data; to center only use the option scale=FALSE, to scale only center=FALSE (by default center=TRUE, scale=TRUE)

setdiff(x, y) 'set' function

setequal(x, y) 'set' function

sin

sum(x) Sum of the elements of x

tan

union(x, y) 'set' function

20 Matrices

%*% Matrix multiplication

colMeans(x) Fast version of col means

colsum(x) Sum of cols for a matrix-like object; colSums(x) is a faster version

diag(x) Diagonal

rowMeans(x) Fast version of row means

rowsum(x) Sum of rows for a matrix-like object; rowSums(x) is a faster version

solve(a) Matrix inverse of a

solve(a, b) Solves a %*% x=b for x

t(x) Transpose

21 Miscellaneous

= Assign

NA Missing data

q() Quit R

setwd('dir') Set R working folder on 'dir'

22 Operators (arithmetic)

- Subtraction

%% Modulo

%/% Integer divide

***** Multiplication

/ Division

Exponentiation

+ Addition

23 Operators (logical)

! Logical negation

!= Not equals

& Elementwise and

& & Control and

$>$ Greater than

$\geq$ Greater than or equal to

$<$ Less than

$\leq$ Less than or equal to

$|$ Elementwise or

$||$ Control or

== Equals

xor Elementwise exclusive or

24 Optimization and model fitting

AIC(fit) Computes the Akaike information criterion or AIC

approx(x, y=) Linearly interpolate given data points; x can be an xy plotting structure

coef(fit) Returns the estimated coefficients (sometimes with their standard-errors)

deviance(fit) Returns the deviance

df.residual(fit) Returns the number of residual degrees of freedom

fitted(fit) Returns the fitted values

glm(formula, family=) Fit generalized linear models, specified by giving a symbolic description of the linear predictor and a description of the error distribution; family is a description of the error distribution and link function to be used in the model; see ?family

lm(formula) Fit linear models; formula is typically of the form response termA + termB + ...; use I(x*y) + I(x^2) for terms made of nonlinear components

loess(formula) Fit a polynomial surface using local fitting

logLik(fit) Computes the logarithm of the likelihood and the number of parameters

nlm(f, p) Minimize function f using a Newton-type algorithm with starting values p

nls(formula) Nonlinear least-squares estimates of the nonlinear model parameters

optim(par, fn, method=c('Nelder', ...), ...) General-purpose optimization; par is initial values, fn is function to optimize (normally minimize)

predict(fit, ...) Predictions from fit based on input data

residuals(fit) Returns the residuals

spline(x, y=) Cubic spline interpolation

25 Processing

aggregate(x, by, FUN) Splits the data frame x into subsets, computes summary statistics for each, and returns the result in a convenient form; by is a list of grouping elements, each as long as the variables in x

apply(x, INDEX, FUN=) A vector or array or list of values obtained by applying a function FUN to margins (INDEX) of x

by(data, INDEX, FUN) Apply FUN to data frame data subsetted by INDEX

lapply(x, FUN) Apply FUN to each element of the list x

merge(a, b) Merge two data frames by common columns or row names

reshape(x, ...) Reshapes a data frame between 'wide' format with repeated measurements in separate columns of the same record and 'long' format with the repeated measurements in separate records; use (direction='wide') or (direction='long')

stack(x, ...) Transform data available as separate columns in a data frame or list into a single column

tapply(x, INDEX, FUN=) Apply FUN to each cell of a ragged array given by x with indexes INDEX

unstack(x, ...) Inverse of stack()

xtabs(a, b, data=x) A contingency table from cross-classifying factors

26 Statistics (mva)

cancor() Canonical correlation

factanal() Factor analysis

kmeans() kmeans cluster analysis

prcomp() Principal components

27 Statistics

anova(fit, ...) Analysis of variance (or deviance) tables for one or more fitted model objects

aov(formula) Analysis of variance model

binom.test() Use help.search('test')

chisq.test(x) Chi-square test on matrix x

cor(x) Correlation matrix of x if it is a matrix or a data frame (1 if x is a vector)

cor(x, y) Linear correlation between x and y, or correlation matrix if they are matrices or data frames

cor.test(a, b) Test correlation

cov(x) Covariance of the elements of x (calculated on n-1); if x is a matrix or a data frame, the variance-covariance matrix is calculated

density(x) Kernel density estimates of x

fisher.test() Fisher exact test

friedman.test() Friedman test

mean(x) Mean of the elements of x

median(x) Median of the elements of x

pairwise.t.test() Use help.search('test')

power.t.test() Use help.search('test')

prop.test() Significance test

quantile(x, probs=) Sample quantiles corresponding to the given probabilities (defaults to 0, 0.25, 0.5, 0.75, 1)

sd(x) Standard deviation of x

t.test() Use help.search('test')

var(x) Variance of the elements of x (calculated on n-1); if x is a matrix or a data frame, the variance-covariance matrix is calculated

var(x, y) Covariance between x and y, or between the columns of x and those of y if they are matrices or data frames

weighted.mean(x, w) Mean of x with weights w

28 Strings

grep(pattern, x) Searches for matches to pattern within x; see ?regex

gsub(pattern, replacement, x) Replacement of matches determined by regular expression matching sub() is the same but only replaces the first occurrence

match(x, table) A vector of the positions of first matches for the elements of x among table

nchar(x) Number of characters

paste(...) Concatenate vectors after converting to character; 'sep=' is the string to separate terms (a single space is the default); 'collapse=' is an optional string to separate 'collapsed' results

pmatch(x, table) Partial matches for the elements of x among table

strsplit(x, split) Split x according to the substring split

substr(x, start, stop) Substrings in a character vector; can also assign, as substr(x, start, stop)=value

tolower(x) Convert to lowercase

toupper(x) Convert to uppercase

x %in% table Idem match(x,table), but returns a logical vector

29 Variable (conversion)

as.array(x) Convert type; for a complete list, use methods(as)

as.character(x) Convert type; for a complete list, use methods(as)

as.complex(x) Convert type; for a complete list, use methods(as)

as.data.frame(x) Convert type; for a complete list, use methods(as)

as.logical(x) Convert type; for a complete list, use methods(as)

as.numeric(x) Convert type; for a complete list, use methods(as)

30 Variable (information)

attr(x,which) Get or set the attribute which of x

attributes(obj) Get or set the list of attributes of obj

class(x) Get or set the class of x; class(x)='myclass'

dim(x) Retrieve or set the dimension of an object; dim(x)=c(3,2)

dimnames(x) Retrieve or set the dimension names of an object

is.array(x) Test for type; for a complete list, use methods(is)

is.character(x) Test for type; for a complete list, use methods(is)

is.complex(x) Test for type; for a complete list, use methods(is)

is.data.frame(x) Test for type; for a complete list, use methods(is)

is.na(x) Test for type; for a complete list, use methods(is)

is.null(x) Test for type; for a complete list, use methods(is)

is.numeric(x) Test for type; for a complete list, use methods(is)

length(x) Number of elements in x

ncol(x) Number of columns; NCOL(x) is the same but treats a vector as a onerow matrix

nrow(x) Number of rows; NROW(x) is the same but treats a vector as a onerow matrix

str(object) Print useful information about object

31 Variable (managing)

ls() Show objects in the search path; specify pat='pat' to search on a pattern

rm(object) Remove object

unclass(x) Remove the class attribute of x

SessionInfo

Windows XP (build 2600) Service Pack 3

- R version 2.10.0 (2009-10-26), `i386-pc-mingw32`
- Locale: `LC_COLLATE=Slovenian_Slovenia.1250`, `LC_CTYPE=Slovenian_Slovenia.1250`, `LC_MONETARY=Slovenian_Slovenia.1250`, `LC_NUMERIC=C`, `LC_TIME=Slovenian_Slovenia.1250`
- Base packages: base, datasets, graphics, grDevices, methods, splines, stats, utils
- Other packages: Hmisc 3.7-0, patchDVI 1.5, survival 2.35-8, XML 2.6-0
- Loaded via a namespace (and not attached): cluster 1.12.1, grid 2.10.0, lattice 0.18-3