

R Functions for Modern Regression

Robert Andersen
McMaster University

July 2003

Below are some of the **R** function used in my Regression III (modern regression) course at the ICPSR Summer Program in Quantitative Methods, University of Michigan. It is not an exhaustive list, but it should be helpful as reference source.

1 Distributions

<i>Package</i>	<i>Function</i>	<i>Description</i>
base	hist(x)	Histogram
	persp(x,y,z)	3D perspective plot
	boxplot(x)	Boxplot
sm	sm.density(x)	Density smooth of histogram
car	qq.plot (variable)	Quantile-comparison plot
	box.cox(var)	Box-Cox transformation of a variable
	scatterplot(x,y)	Enhanced scatterplots
djmrgl	hist3d(x,y)	Dynamic 3D histogram
	persp3d(x,y,z)	Dynamic 3D perspective plot

djmrgl is not on CRAN, but it is available
at Duncan Murdoch's website:
<<http://www.stats.uwo.ca/faculty/murdoch/software>>

2 Regression Diagnostics

<i>Package</i>	<i>Function</i>	<i>Description</i>
base	plot(model)	Basic diagnostics plots
car	cr.plots(model)	Partial residual plots
	av.plots(model)	Partial regression plots
	hatvalues (model)	Hat values
	outlier.test (model)	Test for largest residual
	df.betas(model)	DfBetas measure of influence
	cookd(model)	Cook's D measure of influence
	rstudent(model)	Studentized residuals
	vif(model)	VIF or GVIF for each term in the model

3 Robust Regression

<i>Package</i>	<i>Function</i>	<i>Description</i>
MASS	rlm (y~x)	M-Estimation
lqs	ltsreg (y~x)	Least-Trimmed squares
	lms(y~x)	Least-Median regression

4 Bootstrapping

<i>Package</i>	<i>Function</i>
boot	boot (data, statistic, R)
	plot (boot.object)
	boot.ci (boot.object, index, type = c("norm", "perc", "bca"))
	jack.after.boot (boot.object, index)

5 Generalized Linear Models

<i>Package</i>	<i>Function</i>	<i>Description</i>
MASS	glm (y~x, family)	Most GLMs except multinomial and ordered
	polr (y~x)	Proportional odds logit model
nnet	multinom (y~x)	Multinomial logit
VGAM	vgam	All glm class models, plus also ordered logit and probit, multinomial logit

VGAM is not on CRAN, but a beta-version is available at Thomas Yee's website:

<http://www.stat.auckland.ac.nz/yee/VGAM/>

6 Nonlinear Relationships *within* Linear Model

<i>Package</i>	<i>Function</i>	<i>Description</i>
car	box.tidwell(model)	Box-Tidwell transformation of X's
MASS	boxcox(model)	Box-Cox transformation of Y
base	lm (y~poly(x, degree))	Orthogonal polynomials
splines	lm (y~bs(x))	Cubic B-splines
	lm (y~ns(x))	Natural cubic splines

7 Nonparametric Regression

<i>Package</i>	<i>Function</i>	<i>Description</i>
base	lowess (x,y, f=.)	Lowess smooth
modreg	loess (y~x, span=., degree=1)	Lowess smooth
locfit	locfit (y~x, alpha=.)	Lowess smooth
pspline	sm.spline(x,y)	Smoothing splines

8 Generalized Additive Models

<i>Package</i>	<i>Function</i>	<i>Description</i>
mgcv	gam (y~s(x1), family)	GAM fit
	plot (model)	Plots additive effects
VGAM	vgam (y~s(x1), family)	Vector GAMs (e.g., GAMs for multinomial, ordered logit)

VGAM is not on CRAN, but a beta-version is available at Thomas Yee's website:

<http://www.stat.auckland.ac.nz/yee/VGAM/>

9 Evaluating Models

<i>Package</i>	<i>Function</i>	<i>Description</i>
base	anova(model)	Type I F-test
car	Anova(model)	Type II, III F-tests
relimp	relimp(model, set1, set2)	Relative importance of sets of predictors
base	deviance(model)	Deviance of model
leaps	subset	Subsets (uses AIC, BIC, Mallow's C_p)
car	regsubsets	Use with subset

10 Displaying Effects

<i>Package</i>	<i>Function</i>	<i>Description</i>
effects	all.effects (model)	Table of fitted values
	effect('term', model)	Table of fitted values
	plot (all.effects (model))	Effect display
	plot(effect('term', model))	Effect display
qvcalc	qvcalc(model, 'factor')	Quasi standard errors dummy variables
	plot(qvcalc(model, 'factor'))	Plots Quasi SEs

11 Other Functions

<i>Package</i>	<i>Function</i>	
nlme	lme(y~x)	Mixed models (Contextual effect models; Panel data)
MASS	lm.ridge	Ridge regression
nls	nls (formula, start values)	Nonlinear least squares