Introduction To R Workshop
Selected R Commands
(See Dalgaard, 2008, Appendix C)

ELEMENTARY

Commands

List Objects in Workspace	**ls()**, or **objects()**
Delete object	**rm()**
Search Path	**search()**

Variable Names

Combinations of letters, digits and period. Don’t start with number or period.

Assignments

Assign value to variable	**<– or =**
Assignment “to the right”	**–>**
Global assignment (in functions)	**<<–**

OPERATORS

For Assignment

Addition	**+**
Subtraction, sign	**–**
Multiplication	*****
Division	**/**
Raise to a power	**^**
Integer Division (Quotient)	**%/%**
Remainder From Integer Division	**%%**

Logical, Relational

Equal to	**==**
Not Equal to	**!=**
Less than	**<**
Greater than	**>**
Less than or Equal to	**<=**
Greater than or Equal to	**>=**
Missing x?	**is.na(x)**
Logical AND (Element)	**&**
Logical AND Operator	**&&**
Logical OR (Element)	**|**
Logical OR Operator	**||**
Logical Not	**!**

Introduction To R Workshop
Selected R Commands
(See Dalgaard, 2008, Appendix C)

VECTORS, DATA TYPES

Generating

Create vector of 25 zeroes	**numeric(25)**
Create vector of 25 blanks	**character(25)**
Create vector of 25 Falses	**logical(25)**
Create vector of values	**seq(start#,end#,Increment)**
-4.0, -3.9, -3.8, -3.7, -3.6	**seq(-4.0, -3.6, 0.1)**
Create vector of 1st 10 integers	**1:10 or seq(1,10,1)**
Concatenation vector	**c(comma separated list)**
5 7 9 13 1 2 3 4 5	**c(5,7,9,13,1:5)**
Repetition vector	**rep(value, number of repetitions)**
1 1 1 1 1	**rep(1,5)**
Factor with 3 levels, repeat each level in blocks of 2 up to length 12	**gl(#levels, #block reps, length)**
1 1 2 2 3 3 1 1 2 2 3 3	**gl(3,2,12)**

Coercion

Convert to Numeric	**as.numeric(x)**
Convert to Character	**as.character(x)**
Convert to Logical	**as.logical(x)**
Create factor from vector x	**factor(x)**

DATA FRAMES

Data Frame with 2 named vectors	**data.frame(height = c(165,185), weight = c(90, 65))**
Collect vectors into Data Frame	**data.frame(height, weight) # both predefined**
Select variable **var** in data frame **dfr**	**dfr$var**
Put data frame in search path	**attach(dfr) # a copy of dfr**
Remove data frame from path	**detach()**

Introduction To R Workshop
Selected R Commands
(See Dalgaard, 2008, Appendix C)

NUMERICAL FUNCTIONS

Mathematical

Logarithm of x natural (base e) logarithm	**log(x)**
Base-10 Logarithm	**log10(x)**
Exponential Function e^x	**exp(x)**
Sine Function	**sin(x)**
Cosine Function	**cos(x)**
Tangent Function	**tan(x)**
Arc-sine Function	**asin(x)**
Arc-cosine Function	**acos(x)**
Arc-tangent Function	**atan(x)**
Atan2 Function	**atan2(y, x) # = atan(y/x), when x, y > 0**
Combinations of n objs taken k at a time	**choose(n, k)**
10	**choose(5, 2)**
Factorial	**factorial(x)**
Absolute value	**abs(x)**
Truncated integer	**floor(x)**
Rounded up integer	**ceiling(x)**
(Positive) Square Root	**sqrt(x)** or **x^(0.5)**
Minimum value in vector	**min(x)**
Minimum value over several vectors	**min(x1, x2, x3, ...)**
Maximum value in vector	**max(x)**
Maximum value over several vectors	**max(x1, x2, x3, ...)**
Range of x	**range(x) # = c(min(x), max(x))**
1 9	**range(1:9)**
8	**diff(range(1:9))**
Number of elements in a vector	**length(x)**
Parallel (elementwise) minimum over multiple equally long vectors	**pmin(x1, x2, x3, ...)**
Parallel (elementwise) maximum over multiple equally long vectors	**pmax(x1, x2, x3, ...)**
Number of non-missing elements in a vector	**sum(complete.cases(x))**

Introduction To R Workshop
Selected R Commands
(See Dalgaard, 2008, Appendix C)

NUMERICAL FUNCTIONS (Continued)

Statistical

Arithmetic Mean of x	**mean(x)**
Weighted Arithmetic Mean of x via w	**weighted.mean(x, w)**
5.4	**weighted.mean(1:9, c(rep(.1, 8), .2))**
Trimmed mean, drop top, bottom 0 - 50%	
Drop top, bottom 5% data values	**mean(x, trim=.05)**
Standard Deviation	**sd(x)**
Variance	**var(x)**
Median	**median(x) # = quantile(x, .5)**
Quantile	**quantile(x, p)**
Interquartile Range	**IQR(x)**
Summary: Min, 25%, 50%, 75%, Max	**fivenum(x)**

Mode	**statmod <- function(x) {**
	z <- table(as.vector(x))
	names(z)[z == max(z)]}

"3" "7"	**statmod(c(1, 3, 5, 7, 9, 3, 7))**
Correlation	**cor(x, y)**

INDEXING/SELECTION

nth element	**x[n]**
Subvector containing 1st five elements	**x[1:5]**
Element numbers: n1, n2, n3, n4, n5	**x[c(n1, n2, n3, n4, n5)]**
Selection by logical expression	**x[y <= n1]**
Selection by factor variable	**x[sex == "male"]**
Selection by numeric variable	**i <- c(2,3,5,7,11); x[i]**
Selection by logical variable	**l <- (y <= n1); x[l]**

Matrices And Data Ranges

Fourth row	**m[4,]**
Third Column	**m[, 3]**
Partial Data Frame	**dfr[dfr$var<=n1,]** **subset(dfr, var<=n1)**

Introduction To R Workshop
Selected R Commands
(See Dalgaard, 2008, Appendix C)

DATA INPUT

Built-in data set	**data(name)**
Read from external file	**read.table(“filename”)**
Copy from window and paste into R	**read.table(“clipboard”, header=T)**
Read data from file or url or STDIN	**scan(file=””)** or **scan()**

Arguments to read.table:

First line has variable names	**read.table(“filename”, header=TRUE)**

Comma separated data	**read.table(“filename”, sep=”,”)**
	read.csv(“filename”)

Tab delimited data	**read.delim(“filename”)**
Tab delimited data, “,” is decimal point	**read.delim2(“filename”)**
Decimal point is “,” in data	**read.table(“filename”, dec=”,”)**
Semicolon separated data, decimal point is “,”	**read.csv2(“filename”)**
Missing value is dot (“.”) in data	**read.table(“filename”, na.strings=”.”)**

STATISTICAL DISTRIBUTIONS

Normal Distribution

Normal Density function	**dnorm(x)**
Cumulative Distribution function P{X ≤ x}	**pnorm(x)**
p-quantile x: P{X ≤ x} = p	**qnorm(x)**
n (pseudo) random normally distributed numbers	**rnorm(x)**
Normal Distribution	**pnorm(x, mean, sd)**

Distributions* (also d, q, r prefixes)

Student’s t distribution	**pt(x, df)**
F distribution	**pf(x, n1, n2)**
Chi-squared distribution	**pchisq(x, df)**
Binomial distribution	**pbinom(x, n, p)**
Poisson distribution	**ppois (x, lambda)**
Uniform distribution	**punif (x, min, max)**
Exponential distribution	**pexp (x, rate)**
Hypergeometric distribution	**phyper(x, m, n, k)**

Introduction To R Workshop
Selected R Commands
(See Dalgaard, 2008, Appendix C)

STATISTICAL STANDARD METHODS

Continuous Response

1 and 2 sample t tests	**t.test**
Pairwise comparisons	**pairwise.t.test**
Correlation	**cor.test**
Comparison of 2 variances (F test)	**var.test**
Regression Analysis (linear model)	**lm(y ~ x)**
One way Analysis of Variance	**lm(y ~ f)**
Two way Analysis of Variance	**lm(y ~ f1 + f2)**
Analysis of Covariance	**lm(y ~ f + x)**
Multiple Regression Analysis (linear)	**lm(y ~ x1 + x2 + x3)**
Bartlett test (k variances)	**bartlett.test**

Nonparametric

1 and 2 sample Wilcoxon tests	**wilcox.test**
Kruskal-Wallis Test	**kruskal.test**
Friedman’s 2 way Analysis of Variance	**friedman.test**
Kendall’s Tau (τ) Test	**cor.test (method = “kendall”)**
Spearman’s Rho (ρ) Test	**cor.test (method = “spearman”)**

Discrete Response

Binomial Test (Including Sign test)	**binom.test**
Comparison of Proportions Test	**prop.test**
Confidence Intervals of Proportions	**prop.test(x, n, conf.level=.95, conf.int=TRUE)**
Trend test in Relative Proportions	**prop.trend.test**
Fisher Exact Test in Small Tables	**fisher.test**
Chi Squared test	**chisq.test**
Logistic Regression	**glm(y ~ x1 + x2 + x3, binomial)**

MODELS

Model Formulas

Described By	**~**
Additive Effects	**+**
Interaction	**:**
Main Effects + Interaction	*** (a*b = a + b + a:b)**
Remove Intercept	**-1**

Introduction To R Workshop
Selected R Commands

(See Dalgaard, 2008, Appendix C)

Linear, Non-linear, Generalized Linear

Fit model and save result	lm.out <- lm(y ~ x)
Coefficients, etc.	summary(lm.out)
Analysis of Variance table	anova(lm.out)
Fitted Values	fitted(lm.out)
Residuals	resid(lm.out)
Deviance (Residual Sum of Squares)	deviance(lm.out)
Predictions for new data frame	predict(lm.out, newdata)
Logistic regression	glm(y ~ x, binomial)
Poisson regression	glm(y ~ x, poisson)
Non-linear regression	nls(y ~ exp(-b*x), start=c(a=5, b=.2))

Diagnostics

Studentized residuals	rstudent(lm.out)
Change in β if obs. removed	dfbetas(lm.out)
Change in fit if obs. removed	dffits(lm.out)

GRAPHICS

Standard Plots

Scatterplot (and more)	plot()
Histogram	hist()
Box and whiskers plot	boxplot()
Strip Chart	stripchart()
Bar Chart	barplot()
Dot Diagram	dotchart()
Pie Chart (Cakes, …)	pie()
Interaction Plot	interaction.plot()

Plotting Elements

Lines	lines()
Line (slope-intercept form)	abline()
Points	points()
Line Segments	segments()
Arrows (Note: angle=90 for error bars)	arrows()
Axis	axis()
Frame around Plot	box()
Title above Plot	title()
Text in Plot	text()
Text in Margin	mtext()
List of Symbols	legend()

Introduction To R Workshop
Selected R Commands

(See Dalgaard, 2008, Appendix C)

GRAPHICS (Continued)

Graphical Parameters

Plotting Character (Symbol)	**pch**
Multiframe (Several plots on one)	**mfrow, mfcol**
Plot limits	**xlim, ylim**
Line type/width	**lty, lwd**
Color	**col**
Character size, line spacing in margins	**cex, mex**

PROGRAMMING

Conditional Execution [with alternative]

```
if (p<0.5)
        print("small fraction") [ else
        print("larger fraction") ]
if (0)
        print("0 means false") [ else
        print("non-zero means true") ]
```

Loop over a list

```
for (i in 1:10)
        print (i)
```

Loop

```
i <- 1
while (i < 10)  {
        print(i)
        i <- i + 1 }
```

User defined function (Definition)

```
f <- function(a, b, doit=TRUE) {
     if (doit)
        a + b
     else
        0
```

Use after defining

f(2, 4, doit = FALSE) or
f(2, 4, doit=TRUE)