

Elementare Befehle

<code>ls()</code> oder <code>objects()</code>	Auflistung der Objekte im Workspace
<code>rm(object)</code>	<code>object</code> wird aus dem Workspace entfernt
<code>help(funktion)</code> oder <code>?funktion</code>	Zeigt die Hilfeseite einer Funktion
<code>args(funktion)</code>	Zeigt die Argumente einer Funktion
<code>example(funktion)</code>	Zeigt Beispielaufrufe einer Funktion

Variablennamen

Variablen werden aus Buchstaben (Groß- und Kleinschreibung beachten!), Ziffern und dem Punkt „.“ gebildet, sollten aber nicht mit Ziffern oder Punkt beginnen.

Zuweisungen

<code>&lt;-</code>	Übergabe eines Werts an eine Variable
<code>-&gt;</code>	Übergabe eines Werts an eine Variable nach rechts

Arithmetische Operationen

<code>+</code>	Addition
<code>-</code>	Subtraktion
<code>*</code>	Multiplikation
<code>/</code>	Division
<code>^</code>	Potenzierung
<code>%/%</code>	Division mit Rest
<code>%%</code>	Rest einer Division

Logische Operationen

<code>==</code>	gleich
<code>!=</code>	ungleich
<code>&lt;</code>	kleiner
<code>&gt;</code>	größer
<code>&lt;=</code>	kleiner gleich
<code>&gt;=</code>	größer gleich
<code>&amp;</code>	logisches UND
<code> </code>	logisches ODER
<code>!</code>	logisches NICHT

Generierung von Vektoren

<code>c(7,2,13)</code>	generiert den Vektor 7 2 13
<code>1:4</code>	generiert den Vektor 1 2 3 4
<code>4:1</code>	generiert den Vektor 4 3 2 1
<code>seq(0,1,0.1)</code>	generiert den Vektor 0 0.1 ... 0.9 1
<code>rep(c(2,7),2)</code>	erzeugt den Vektor 2 7 2 7
<code>rep(c(2,7),2:3)</code>	erzeugt den Vektor 2 2 7 7 7
<code>numeric(10)</code>	ein Vektor mit 10 Nullen
<code>character(10)</code>	10 × " "
<code>logical(10)</code>	10 × FALSE

Umwandlung

<code>as.character(x)</code>	wandelt in Zeichenketten-Variable um
<code>as.numeric(x)</code>	wandelt in numerische Variable um
<code>factor(x)</code>	konvertiert in einen nominalskalierten Vektor

Datentabellen

<code>data.frame(x,y)</code>	erzeugt Datentabelle mit den Spalten <code>x</code> und <code>y</code>
<code>data.frame(age=x,height=y)</code>	Datentabelle mit zwei benannten Spalten <code>x</code> und <code>y</code>
<code>attach(data)</code>	<code>data</code> wird in den Suchpfad kopiert
<code>detach(data)</code>	<code>data</code> wird wieder aus dem Suchpfad gelöscht

Einlesen von Datensätzen

<code>read.table("Pfadname")</code>	liest externen Datensatz ein
-------------------------------------	------------------------------

Gebräuchliche Argumente der read.table-Funktion

<code>header=TRUE</code>	die erste Zeile wird als Zeile der Variablennamen aufgefasst
<code>sep=", "</code>	Daten werden durch Kommata getrennt
<code>dec=","</code>	Komma ist Dezimalkomma (üblich: Punkt)
<code>na.strings=". "</code>	Punkte für fehlende Einträge

Varianten von read.table

<code>read.csv("Pfadname")</code>	liest durch Kommata getrennte Spalten
<code>read.csv2("Pfadname")</code>	liest durch Semikola getrennte Spalten, Komma als Dezimalkomma
<code>read.delim("Pfadname")</code>	liest Tab-getrennte Spalten
<code>read.delim2("Pfadname")</code>	liest Tab-getrennte Spalten, Komma als Dezimalkomma

Indizierung von Vektoren

<code>x[i]</code>	gibt die <i>i</i> -te Komponente des Vektors <code>x</code> aus
<code>x[-i]</code>	gibt alle bis auf die <i>i</i> -te Komponente des Vektors <code>x</code> aus
<code>x[1:5]</code>	gibt die ersten 5 Komponenten von <code>x</code> aus
<code>x[c(2,3,5)]</code>	gibt die 2., 3. und 5. Komponente des Vektors <code>x</code> aus
<code>x[y&lt;=30]</code>	gibt den Vektor der Komponenten <code>x<sub>i</sub></code> aus, für die <code>y<sub>i</sub> ≤ 30</code> ist
<code>which(y&lt;=30)</code>	gibt die Positionen der Komponenten <code>y<sub>i</sub></code> aus, für die <code>y<sub>i</sub> ≤ 30</code> ist
<code>x[sex="male"]</code>	Selektion anhand eines qualitativen Merkmals

Indizierung von Datentabellen und Matrizen

<code>dfr[4,3]</code>	gibt Eintrag aus der 4. Zeile/3. Spalte aus
<code>dfr[4,]</code>	gibt Vektor der 4. Zeile aus
<code>dfr[,3]</code>	gibt Vektor der 3. Spalte aus
<code>dfr\$body</code>	gibt Spalte <code>body</code> der Datentabelle <code>dfr</code> aus
<code>subset(dfr,body&lt;=70)</code>	gibt Teiltabelle der Datentabelle aus, die in der Spalte <code>body</code> einen Wert <code>≤ 70</code> haben

Elementare numerische Funktionen

<code>abs(x)</code>	Betrag von <code>x</code>
<code>acos(x)</code>	Arkuskosinus von <code>x</code>
<code>asin(x)</code>	Arkussinus von <code>x</code>
<code>atan(x)</code>	Arkustangens von <code>x</code>
<code>cos(x)</code>	Kosinus von <code>x</code>
<code>choose(n,k)</code>	Binomialkoeffizient $\binom{n}{k}$
<code>exp(x)</code>	$e^x$
<code>factorial(n)</code>	Fakultät von <code>n</code>
<code>log(x)</code>	natürlicher Logarithmus von <code>x</code>
<code>log10(x)</code>	10er Logarithmus von <code>x</code>
<code>round(x,3)</code>	Rundet <code>x</code> auf 3 Stellen
<code>sin(x)</code>	Sinus von <code>x</code>
<code>tan(x)</code>	Tangens von <code>x</code>

Funktionen für Vektoren

<code>all(x,cond)</code>	Gibt TRUE, wenn alle Komponenten die Bedingung <code>cond</code> erfüllen
<code>any(x,cond)</code>	Gibt TRUE, wenn eine Komponente die Bedingung <code>cond</code> erfüllt
<code>cumsum(x)</code>	Vektor der kumulierten Summen der Komponenten von <code>x</code>
<code>cumprod(x)</code>	Vektor der kumulierten Produkte der Komponenten von <code>x</code>
<code>length(x)</code>	Anzahl der Komponenten des Vektors <code>x</code>
<code>min(x)</code>	Minimum der Komponenten des Vektors <code>x</code>
<code>min(x<sub>1</sub>, ..., x<sub>n</sub>)</code>	Min. der Komponenten der Vektoren <code>x<sub>1</sub>, ..., x<sub>n</sub></code>
<code>max(x)</code>	Maximum der Komponenten des Vektors <code>x</code>
<code>pmin(x<sub>1</sub>, ..., x<sub>n</sub>)</code>	wie <code>c(min(c(x<sub>1</sub>[1], ..., x<sub>n</sub>[1])), ... )</code>
<code>pmax(x)</code>	analog zu <code>pmin</code>
<code>prod(x)</code>	Produkt der Komponenten des Vektors <code>x</code>
<code>range(x)</code>	wie <code>c(min(x),max(x))</code>
<code>sort(x)</code>	Aufsteigende Sortierung des Vektors <code>x</code>
<code>sum(x)</code>	Summe der Komponenten des Vektors <code>x</code>

Funktionen für Matrizen

<code>t(M)</code>	Transponiert die Matrix <code>M</code>
<code>M %*% N</code>	Matrizenmultiplikation
<code>det(M)</code>	Determinante von <code>M</code>
<code>solve(M,b)</code>	löst das LGS <code>Mx=b</code>
<code>colSums(M)</code>	Spaltensummen von <code>M</code>
<code>rowSums(M)</code>	Zeilensummen von <code>M</code>
<code>colMeans(M)</code>	Mittelwerte der Spalten von <code>M</code>
<code>rowMeans(M)</code>	Mittelwerte der Zeilen von <code>M</code>

Funktionen aus der Statistik

<code>mean(x)</code>	Durchschnitt von <code>x</code> : $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
<code>var(x)</code>	Stichprobenvarianz von <code>x</code> : $\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$
<code>sd(x)</code>	Standardabweichung, wie <code>sqrt(var(x))</code>
<code>cor(x,y)</code>	Korrelationskoeffizientenmatrix von <code>x</code> und <code>y</code>
<code>quantile(x,p)</code>	p-Quantil von <code>x</code>
<code>median(x)</code>	Median von <code>x</code> , wie <code>quantile(x,0.5)</code>
<code>summary(x)</code>	Zusammenstellung wichtiger Statistiken von <code>x</code>
<code>cut(x,breaks)</code>	teilt den Wertebereich von <code>x</code> in Intervalle mit den in <code>breaks</code> angegebenen Intervallgrenzen und kodiert die Werte von <code>x</code> entsprechend der Klasse, in die sie fallen
<code>table(x)</code>	Kontingenztafel von <code>x</code>
<code>prop.table(x,margin)</code>	bedingte Häufigkeiten gegeben die Zeilen, falls <code>margin=1</code> bedingen unter Spalten mit <code>margin=2</code>
<code>order(x)</code>	Ordnungsstatistik angewandt auf <code>x</code>
<code>rank(x)</code>	gibt die Rangwerte des numerischen Vektors <code>x</code> wieder, bei Bindungen oder Ties werden mittlere Ränge berechnet
<code>lm(y~x)</code>	bestimmt Regressionsparameter der lin. Regression der abhängigen Variablen <code>y</code> und der unabh. <code>x</code>

Programmierung

<code>if(cond){commands}</code>	Ausführung der Befehle <code>commands</code> , wenn der Wert der Formel <code>cond</code> TRUE ist
<code>else{commands}</code>	alternative Anweisungen (im Anschluss an if-Befehl)
<code>ifelse(cond, x, y)</code>	ist <code>cond</code> TRUE wird <code>x</code> zurückgegeben, sonst <code>y</code>
<code>for(i in x){commands}</code>	Schleife über die Werte des Vektors <code>x</code>
<code>while(cond){commands}</code>	Wiederholung der Befehle <code>commands</code> solange der Wert der Formel <code>cond</code> TRUE ist
<code>repeat {commands}</code>	(erstmal) unbegrenzte Wiederholung der Befehle <code>commands</code>
<code>if(cond) break</code>	wenn der Wert der Formel <code>cond</code> TRUE ist, wird die aktuelle repeat-Schleife abgebrochen
<code># Kommentar</code>	alle Zeichen hinter <code>#</code> werden nicht aufgeführt

Normalverteilung: Dichte, Verteilungsfunktion, Quantile und Zufallszahlen

<code>dnorm(x)</code>	Dichte der Standardnormalverteilung in <code>x</code>
<code>pnorm(x)</code>	Verteilungsfunktion der Standardnormalerteilung in <code>x</code>
<code>qnorm(p)</code>	p-Quantil der Standardnormalverteilung, d. h., <code>qnorm(p)</code> ist das <code>x</code> mit $\mathcal{N}(0,1)((-\infty, x]) = p$
<code>rnorm(n)</code>	erzeugt <code>n</code> (Pseudo-) Zufallszahlen gemäß der $\mathcal{N}(0,1)$ -Vtlg.

Andere Verteilungsfunktionen

<code>pbinom(x,n,p)</code>	Binomialverteilung mit Parametern <b>n</b> und <b>p</b>
<code>pgeom(x,p)</code>	Geometrische-Verteilung mit Parameter <b>p</b>
<code>phyper(x,m,n,k)</code>	Hypergeometrische-Verteilung
<code>pnbinom(x,size,prob)</code>	negative Binomialerteilung
<code>ppois(x,lambda)</code>	Poisson( $\lambda$ )-Verteilung
<code>punif(x,a,b)</code>	Gleichverteilung auf <b>(a,b)</b>
<code>pexp(x,rate)</code>	Exponentialverteilung
<code>pbeta(x,a,b)</code>	$\beta$ ( <b>a,b</b> )-Verteilung
<code>pchisq(x,df)</code>	$\chi^2$ -Verteilung
<code>pf(x,df1,df2)</code>	<i>F</i> -Verteilung
<code>pgamma(x,a,b)</code>	$\Gamma$ ( <b>a,b</b> )-Verteilung
<code>pweibull(x,a,b)</code>	Weibull-Verteilung mit Parametern <b>a</b> und <b>b</b>
<code>pcauchy(x,a,b)</code>	Cauchy-Verteilung mit Parametern <b>a</b> und <b>b</b>
<code>plnorm(x,mean,sd)</code>	Lognormalverteilung
<code>plogis(x,location,scale)</code>	Logistische Verteilung
<code>pnorm(x,mean,sd)</code>	Normalverteilung
<code>psignrank(x,n)</code>	Verteilung der Wilcoxon'schen Vorzeichen-Statistik
<code>pt(x,df)</code>	Student'sche <i>t</i> -Vtlg. ( <b>df</b> für <i>degrees of freedom</i> )
<code>pwilcox(x,m,n)</code>	Vtlg. der Wilcoxon'schen Rangsummenstatistik für Samples der Größe <b>m</b> und <b>n</b>

Ersetzt man das **p** in den Befehlen durch **d** oder **q**, so erhält man die Dichte oder ein Quantil der entsprechenden Verteilung. Ersetzt man das **p** durch ein **r**, so erhält man eine Randomisierung (auch: *Sample*) gemäß der jeweiligen Verteilung.

Testen in R

<code>binom.test</code>	Exakter Binomialtest
<code>chisq.test</code>	Chiquadrat-Anpassungstest
<code>ks.test</code>	Kolmogoroff-Smirnoff-Test
<code>shapiro.test</code>	Shapiro-Wilk-Normalverteilungstest
<code>t.test</code>	Student'scher <i>t</i> -Test und Welch'scher <i>t</i> -Test
<code>wilcox.test</code>	Wilcoxon-Vorzeichen-Rangtest und Wilcoxon-Rangsummentest

Standard-Plots

<code>barplot()</code>	Säulendiagramm
<code>boxplot()</code>	Boxplot
<code>curve(f(x),a,b)</code>	zeichnet den Graphen der Funktion <i>f</i> auf <b>[a,b]</b>
<code>hist()</code>	Histogramm
<code>interaction.plot()</code>	Profildiagramm
<code>mosaicplot()</code>	Mosaikplot
<code>plot()</code>	Standardplot
<code>plot(ecdf(x))</code>	Empirische Verteilungsfunktion des Vektors <b>x</b>
<code>pie()</code>	Kreisdiagramm
<code>qqplot()</code>	QQ-Plot
<code>qqnorm()</code>	NQ-Plot

Zusätzliche Plot-Elemente

<code>abline(a,b)</code>	fügt Gerade mit Achsenabschnitt <b>a</b> und Steigung <b>b</b> in Plot ein
<code>abline(h=u)</code>	wie oben; jedoch Horizontale in <b>u</b>
<code>abline(v=u)</code>	wie oben; jedoch Vertikale in <b>u</b>
<code>grid()</code>	zeichnet Gitterlinien in die Graphik
<code>box()</code>	zeichnet einen Rahmen um die Graphik
<code>legend()</code>	zeichnet eine Legende
<code>lines()</code>	zusätzliche Linienelemente für einen Plot
<code>points()</code>	zusätzliche Punkte für einen Plot
<code>title()</code>	erstellt Titel (und Untertitel) für den Plot

Graphikparameter

<code>expression</code>	mit <b>xlab</b> und <b>ylab</b> , erlaubt das Einsetzen von mathematischen Termen (z. B. mit Index)
<code>lty, lwd</code>	Linientyp bzw. Linienstärke
<code>mfrow, mfcoll</code>	mehrere Graphiken in einem Bild ( <b>multiframe</b> )
<code>pch</code>	Einstellung des <b>plotting characters</b>
<code>xlim=c(0,9)</code>	die Graphik zeigt den <b>x</b> -Achsenabschnitt von 0 bis 9
<code>ylim=c(0,9)</code>	wie <b>xlim</b> für die <b>y</b> -Achse

Speichern von Graphiken

<code>pdf("Pfadname")</code>	Speichert Graphik des folgenden <b>plot</b> -Befehls in eine <i>pdf</i> Datei
<code>jpeg("Pfadname")</code>	Speichert Graphik des folgenden <b>plot</b> -Befehls in eine <i>jpg</i> Datei
<code>png("Pfadname")</code>	Speichert Graphik des folgenden <b>plot</b> -Befehls in eine <i>png</i> Datei
<code>bmp("Pfadname")</code>	Speichert Graphik des folgenden <b>plot</b> -Befehls in eine <i>bmp</i> Datei
<code>tiff("Pfadname")</code>	Speichert Graphik des folgenden <b>plot</b> -Befehls in eine <i>tiff</i> Datei

Im Anschluss des **plot**-Befehls muss **dev.off()** aufgerufen werden.  
Beispiel: `pdf("Pfadname"); plot(ecdf(rnorm(15))); dev.off();`