

Gyakori R parancsok

Oktatási segédlet, v0.5

Összeállította: Abari Kálmán (abari.kalman@gmail.com)

Segítség kérése

`help(téma)`, `?téma` a téma leírása, pl.: `help(t.test)` vagy `?t.test`

`help.search("téma")`, `??téma`, `??"téma"` téma előfordulásának keresése

`find("téma")` a témát tartalmazó csomag listázása

`apropos("téma")` objektumok listája, amelyekben a téma előfordul

`help.start()` HTML help indítása

`vignette()` a csomagokhoz mellékelte dokumentációk megtekintése és olvasása, pl.: `vignette(package="HSAUR")`

`example(téma)` a témához tartozó példák futtatása

`demo()` beépített demó-szkriptek futtatása, pl.: `demo(graphics)`

`RSiteSearch("téma")` keresés indítása a böngészőben a <http://search.r-project.org> címről

Csomagok kezelése

`installed.packages()` a telepített csomagok listája

`install.packages("csomagnév")` csomag letöltése, telepítése

`update.packages()` csomagok frissítése

`remove.packages("csomagnév")` csomag eltávolítása a számítógépről

`packageStatus()` telepített csomagjaink állapota

`library(csomagnév)` csomag betöltése, elhelyezése a keresési útb

`detach(package=csomagnév)` csomag eltávolítása a memóriából, törlés az útból

`search()` itt: a betöltött csomagok listája

`help(package=csomagnév)` a csomag rövid leírása

`library(help=csomagnév)` a csomag rövid leírása

`ls("package:csomagnév")` a csomagban lévő objektumok listája

`data(package="csomagnév")` a csomagban lévő adatobjektumok listája

R-környezet parancsai

`options(...)` globális beállítások, pl.: `options(digits=3)` 3 tizedesre kerekít;

`options(continue=" + ")` a sor folytatását jelezzük másképp

`search()` keresési utak (környezetek) listája

`attach(what)`, `detach(name)` adattábla vagy lista elhelyezése a keresési útb

ill. törlése a keresési útból, pl.: `attach(women)`; `detach(women)`, ld. `with()`

`ls(name)`, `objects(name)` a name környezetben lévő objektumok listája

`rm(...,list)` a felsorolt objektum eltávolítása, pl.: `rm(x)`, `rm(list=ls())`

`getwd()`, `setwd(dir)` a munkakönyvtár lekérdezése és beállítása

`dir(path,pattern)` könyvtár tartalmának listázása, pl.: `dir(pattern="x")`

Információ objektumokról

`class(x)` x osztályának lekérdezése és beállítása

`mode(x)` x módja

`length(x)` x elemeinek a száma

`typeof(x)` x típusa

`str(x)` x szerkezete

`attributes(x)` x attribútumainak lekérdezése vagy beállítása

`attr(x, which)` a which attribútum lekérdezése vagy beállítása

`names(x)` x elemeinek a neve

`dim(x)` x dimenzióinak lekérdezése vagy beállítása

`nrow(x)`, `ncol(x)` x sorainak ill. oszlopainak a száma

`dimnames(x)` x dimenziói mentén a nevek lekérdezése vagy beállítása

`rownames(x)` a sorok neveinek lekérdezése vagy beállítása

`colnames(x)` az oszlopok neveinek lekérdezése vagy beállítása

`levels(x)` a faktor szintjeinek lekérdezése vagy beállítása

`nlevels(x)` a faktor szintjeinek száma

`is.na(x)`, `is.null(x)`, `is.nan(x)`, `is.numeric(x)`, `is.logical(x)`, `is.character(x)`, `is.vector(x)`, `is.factor(x)`, `is.ordered()`, `is.list(x)`, `is.matrix(x)`, `is.data.frame(x)` `stb.` az x objektum tesztelése a függvényneveben szereplő típusra

Operátorok

`<-`, `<<-`, `->`, `->>`, = értékadás operátor, pl.: `sulyok<-c(72,81)`

`+`, `-`, `*`, `/` a négy alpművelet aritmetikai operátora, pl.: `(3-2)/sulyok`

`^`, `**` a hatványozás operátorai, pl. `1024 a 2^10` vagy `2**10` is

`%%`, `%%` maradékos osztás, ill. egész osztás, pl.: `19%%7` az 5, ill. `19%/7` az 2

`%*%` mátrixok szorzása, pl. `m1%*m2`; belső szorzatot ad ha `m1` és `m2` vektorok

`<`, `<=`, `>`, `>=`, `=`, `!=`, `%in%` logikai operátorok, pl.: `a 2<=3` igaz (TRUE)

`!`, `&`, `&&`, `|`, `||` logikai összekötők: a NEM az ÉS és a VAGY; a `&&` és a `||` csak az első elemeket figyel, a többi elemenként hajtja végre a műveletet

`[`, `[[`, `$` vektor, faktor, mátrix, tömb, lista, stb. indexelése, ld. *Indexelés*

`+`, `-` előjel operátorok (unáris), pl.: `-3.4`, `+2`, `-c(1,2)`

`:` (kettőspont) számtani sorozat elemei 1 lépésközzel, pl.: `2:6`

`~` (tilde) formula építése, ld. *Statistikai formulák* rész

Input/output

`scan(what)`, `fix(x)`, `edit(name)` adatok begépelése

`readClipboard()` olvasás vágóasztalról

`cat(...,file,sep)` objektumok kiírása konzolra vagy állományba

`print(x,digits,...)` általános objektumkiíró (digits: tizedesjegyek száma)

`format(x,...)` objektumok formázott kiírása, továbbá: `formatC()`, `sprintf()`

`sink(file)` az output átirányítása file-ba a következő `sink(file=NULL)` hívásig

`read.table()`, `read.csv()`, `read.csv2()`, `read.delim()`, `read.delim2()`

adattábla beolvasása tagolt szöveges állományból (szokásos

paraméterek: file, header, sep, dec, row.names, na.strings)

`read.fwf(file,widths)` fix oszlopszéles állomány beolvasása

`write.table(x)`, `write.csv(x)`, `write.csv2(x)` adattábla vagy mátrix

kiírása tagolt szöveges állományba (szokásos paraméterek: file, sep, dec, row.names, col.names, na)

`load(file)`, `save(x,file)`, `save.image(file)` R objektumok betöltése és kiírása (a bináris `*.RData` és a korábbi `*.rda` állományok kezelése)

`dget(file)`, `dput(x, file)` objektumok szöveges állományon alapuló beolvasása és kiírása

`loadhistory(file)` parancsok történetének betöltése

`savehistory(file)` parancsok történetének kiírása

További lehetőségek pl.: `help(package=foreign)`, `help(package=RODBC)`

Adatobjektumok létrehozása

`c(...)` az argumentumok összefűzése, pl.: `c(1,2:6,7)`

`:` (kettőspont) számtani sorozat elemei 1 lépésközzel, pl.: `2:6`

`seq(from, to, by, length.out, along.with)` általános számtani sorozat elemeit hozhatjuk létre, pl.: `seq(from=3,to=27,by=1.5)`

`rep(x, times, each)` vektorlétrehozás ismétléssel, pl.: `rep(c(1,2),c(3,4))`

`vector(mode, length, numeric(length), character(length), logical(length))` vektor létrehozása az alapvető attribútumok alapján pl.: `vector("numeric",20)`, `logical(10)`

`factor(x)`, `ordered(x)` faktor ill. rendezett faktor létrehozása

`gl(n,k,length,labels)` faktor létrehozása minta alapján pl.: `gl(3,2,7)`

`matrix(x,ncol,nrow,byrow=F)` mátrix létrehozása x vektor alapján, pl.:

`matrix(1:12,nrow=3,byrow=T)` 3x4-es mátrix sorfolytonosan feltöltve

`array(data,dim)` tömb létrehozása a data vektorból, pl.: `array(1,c(2,3,6))`

`list(...)` lista létrehozása, pl.: `list(id=0:9,letszam=10,korosztaaly="idős")`

`data.frame(...)` adattábla létrehozása, pl.: `data.frame(nev=x,suly=y)`, ahol x és y azonos hosszúságú karakteres ill. numerikus vektorok

`rbind(...)`, `cbind(...)` az argumentumok (vektor, mátrix és adattábla), mint sorok ill. oszlopok összefűzése, pl.: `rbind(1:10,1)`

`expand.grid(...)` a megadott vektor v. faktor minden érték kombinációja a létrejövő adattábla egy sora lesz, pl.: `expand.grid(x=1:2,y=3:4)`

Indexelés

☐ Ha x vektor vagy faktor:

`x[n]` az n. elem

`x[-n]` az összes elem, kivéve az n.

`x[c(1,4,3)]` az 1., 4. és 3. elem

`x[-c(1,4,3)]` az összes elem, kivéve az 1., 4. és 3.

`x["név"]` a "név" nevű elem

`x[c("név1","név2")]` a "név1" és "név2" nevű elem

`x[c(T,F)]` x páratlan sorszámú elemei

`x[x > 3]` a 3-nál nagyobb elemek x-ből

☐ Ha x lista:

`x[n]` egyelemű lista, amelyik az x lista n. elemét tartalmazza

`x[[n]]` az x lista n. eleme

`x[["név"]]` az x lista "név" nevű eleme

`x$név` az x lista "név" nevű eleme

☐ Ha x mátrix:

`x[i,j]` az i. sor j. oszlopában lévő elem

`x[i,]` az i. sor minden eleme

`x[,j]` a j. oszlop minden eleme

továbbá a vektornál felsorolt összes lehetőség az i és j helyén.

☐ Ha x adattábla (dataframe):

a listánál és a mátrixnál felsorolt összes lehetőség.

Adatkezelés

head(x,n=6), **tail(x,n=6)** x adattábla vagy mátrix első ill. utolsó sorai
rev(x) x elemei sorrendjének megfordítása, pl.: `rev(1:10)`
unique(x) az x (vektor v. adattábla) duplikált elemeit v. sorait távolítja el
all.equal(x,y) x és y objektum teljes egyezőségét teszteli
with(data,expr) a data adattábla környezetében `expr` kiértékelése

- ▢ Hiányzó adatok (NA) kezelése:

complete.cases(...) hiányzó adatok jelzése logikai értékekkel vektorban, mátrixban vagy adattáblában, ld. még `is.na()`

na.omit(x) a hiányzó adatokat tartalmazó értékek v. sorok eltávolítása

- ▢ Továbbá számos függvény argumentumában:

na.rm statisztikai fv-ekben (`mean()` stb.) gyakori; ha TRUE kihagyja az NA -kat
na.strings a `read.table()` fv-ekben az NA jelölése a szöveges állományban
na.last a rendező fv-ekben használt argumentum; ha értéke TRUE hátra, ha FALSE előre és ha NA kihagyásra kerülnek a hiányzó adatok

na.action különböző modell fv-ekben (`lm()`, `nls()` stb.) az NA-k kezelése:

- "na.fail", "na.omit", "na.exclude"; ld. `options("na.action")` értékét
- ▢ Szűrés és rendezés (szűréshez ld. még az *Indexelés* részt)

which.min(x), **which.max(x)** az x vektor legkisebb ill. legnagyobb elemének az indexe, pl.: `which.max(c(2:5,3))` kifejezés értéke 4

which(x) a logikai x vektor TRUE értékeinek indexét adja, pl.: `which(x>5)`

subset(x,subset,select) az x adattábla subset logikai feltételnek eleget tevő sorait és a select-ben lévő oszlopait válogatja le

sort(x, decreasing=F) x elemeinek rendezett sorrendjét adja

order(x, decreasing=F) x indexeinek az a permutációja, mely x-et rendezi

- ▢ Adatok konvertálása, átkódolása:

as.numeric(x), **as.character(x)**, **as.logical(x)**, **as.vector(x)**, **as.factor(x)**, **as.ordered()**, **as.list(x)**, **as.matrix(x)**, **as.data.frame(x)** stb. x konvertálása a függvény neve szerinti típusba

cut(x,breaks,labels) az x numerikus vektor faktorrá konvertálása a breaks-ben szereplő intervallumhatárok alapján; szintnevek a labels-ben

recode(var,recodes) a var vektor vagy faktor értékeit változtatjuk meg recodes karakterlánc alapján, pl.: `x<-1:10; recode(x,"3=100; 5:7=100; else=0")` vagy `recode(x,"l0:4='a';c(5,7)='k';8:hi='f'")` Csomag: `car`

transform(adattábla,név=érték) az adattábla régi vagy új oszlopának adunk új értéket, pl.: `data(women); transform(women,mgs=height*2.54)`

within(data,expr) a data adattábla régi és új oszlopainak adunk értéket az `expr`-ben; köztes számítások is, de fontos ezek törlése: `rm()`; ld. `with()`; pl.:`data(women);within(women,{mgs<-height*2.54;i<-1;rm(height,i)}`)

Karaktorsorozatok kezelése

paste(...,sep=" ") karakterek összefűzése a sep elválasztóval
nchar(x) x karaktorsorozat hossza
tolower(x), **toupper(x)** x kisbetűsre ill. nagybetűsre konvertálása
chartr(old,new,x) x-ben az old karaktereit a new karaktereire cseréli
substr(x,start,stop) x sztringnek a numerikus start és stop közötti részre
strsplit(x,split) x sztring feldarabolása split alapján

grep(pattern,x) x-ben pattern reguláris kif-sel egyező részsstring keresése
sub(pattern,replacement,x), **gsub()** x-ben a pattern reguláris kif-sel egyező első részt replacemenet-re cseréli; `gsub()` ua., de mindet cseréli

Matematikai függvények

abs(x), **sign(x)** abszolút érték ill. előjel függvény
sqrt(x) négyzetgyök függvény
sin(x), **cos(x)**, **tan(x)**, **asin(x)**, **acos(x)**, **atan2(x)**, **sinh(x)**, **cosh(x)** stb. trigonometrikus függvények
log(x,base), **log10(x)**, **log2(x)** logaritmus függvények
exp(x) exponenciális függvény, pl.: `e=exp(1)`
sum(x), **prod(x)** x elemeinek összege ill. szorzata
diff(x,lag=1) x vektor szomszédos (vagy távolabbi) elemeinek különbsége
round(x,digits=0) x kerekítése digits tizedes számjegyre
ceiling(x) x-nél nagyobb egészek közül a legkisebb
floor(x) x-nél kisebb egészek közül a legnagyobb
trunc(x) x-hez legközelebbi egész x és 0 között (a tizedes jegyek elhagyása)
choose(n,k) n elem k elemű ismétlés nélküli kombinációinak száma
scale(x,center=T,scale=T) alapbeállításban Z-transzformációt végez
sample(x,size,replace=F) x vektorból size elemű véletlen minta kiválasztása visszatevés nélkül vagy visszatevéssel (`replace=T`)
t(x) x mátrix transzponálása (sorok és oszlopok felcserélése)
det(x) x mátrix determinánsa
eigen(x) x mátrix sajátvektorainak és sajátértékeinek számítása; listát ad

Eloszlások

Mindegyik eloszláshoz négy függvény tartozik, amelyek csak az egybetűs előtagban (d, p, q és r) különböznek. A d sűrűség, a p eloszlás, a q kvantilis és az a r véletlen függvényeket jelöl. Az eloszlásra jellemző közös argumentumok az egyes esetekben a következőképp egészülnek ki: *df*függv(x,...), *pf*függv(q,...), *qf*függv(p,...), *rf*függv(n,...). A véletlen számokat generáló példányok:

rbeta(n, shape1, shape2) béta eloszlás
rbinom(n, size, prob) binomiális eloszlás
rcauchy(n, location=0, scale=1) Cauchy eloszlás
rchisq(n, df) χ2 eloszlás
rexp(n, rate=1) exponenciális eloszlás
rf(n,df1,df2) Fisher F-eloszlás
rgamma(n, shape, rate=1) Gamma eloszlás
rgeom(n, prob) geometriai eloszlás
rhyper(nn, m, n, k) hipergeometrikus eloszlás
rlogis(n, location=0, scale=1) logisztikus eloszlás
rlnorm(n, meanlog=0, sdlog=1) normális eloszlás
rnbinom(n, size, prob) negatív binomiális eloszlás
rnorm(n, mean=0, sd=1) normális eloszlás
rpois(n, lambda) Poisson-eloszlás
rwilcox(nn, n, m), **rsignrank(nn, n)** Wilcoxon-eloszlás
rt(n, df) t-eloszlás
runif(n, min=0, max=1) egyenletes eloszlás

rweibull(n, shape, scale=1) Weibull-eloszlás

Leíró statisztika 1. (mutatók, táblázatok)

summary(x) itt: x leíró statisztikai mutatói, x tipikusan numerikus vektor, faktor vagy adattábla, pl.: `summary(d)`, `summary(d$su1y)`, ahol d adattábla
mean(x), **weighted.mean(x,w)** x átlaga, ill. w-vel súlyozott átlaga
median(x), **quantile(x, probs)** x mediánja ill. probs-ban lévő kvantilisei
sd(x), **var(x)** x szórása ill. varianciája
cor(x,y,method), **cov(x,y,method)** korreláció ill. kovariancia x és y között; a method lehetséges értékei: "pearson", "kendall", "spearman"
IQR(x), **mad(x)** interkvartilis terjedelem, ill. abszolút eltérés a mediántól
range(x), **min(x)**, **max(x)** x terjedelme, ill. minimuma és maximuma
kurtosis(x), **skewness(x)** x csúcossága és ferdesége; Csomag: `e1071`
fivenum(x), **boxplot.stats(x)** a Tukey-féle 5 szám, ill. a dobozdiagram számszerűsített adatai
rank(x,ties.method) x rangszámainak meghatározása; kapcsolt rang is
table(...) gyakorisági táblázat készítése faktor argumentum(ok)ból
ftable(...) többdimenziós gyakorisági táblázat megjelenítése
xtabs(formula,data) gyakorisági táblázat formula megadásával
prop.table(x,margin) relatív gyakoriság az egész táblázatra, vagy soronkénti (`margin=1`) ill. oszloponkénti (`margin=2`) relatív gyakoriság
cumsum(x), **cumprod(x)**, **cummax(x)**, **cummin(x)** x vektorra kumulatív összeget, szorzatot, maximumot és minimumot tartalmazó vektort ad
pmax(...), **pmin(...)** párhuzamos maximum és minimum számítása
rowSums(x), **colSums(x)**, **rowMeans(x)**, **colMeans(x)** mátrix vagy numerikus adattábla esetén a sorok ill. oszlopok összege ill. átlaga

Leíró statisztika 2. (grafika)

barplot(height,beside=F,horiz=F) oszlopdiaagram rajzolása (pl. gyakorisági táblázatra); a height numerikus vektor vagy mátrix lehet
stripchart(x,method="overplot") 1 dimenziós pontdiaagram rajzolása; a megjelenítés módja lehet "overplot", "stack" és "jitter"; az x numerikus vektor, lista vagy formula is lehet
dotchart(x) 1 dimenziós ún. Cleveland pontdiaagram rajzolása; az x numerikus vektor vagy mátrix lehet
hist(x,right=T,freq=T,breaks) hisztogram rajzolása x vektorra; alapért. a jobbról zárt intervallum (right) és a gyakoriság megjelenítése (freq)
boxplot(x,range=1.5,notch=F) dobozdiagram megjelenítése; az x numerikus vektor, lista vagy formula is lehet
plot(x), **plot(x,y)**, **plot(y~x)** általános rajzparancs; 2D pontdiaagram
qqplot(x,y), **qqnorm(y)**, **qqline(y)** QQ-ábra rajzolása
pairs(x) x mátrix vagy adattábla oszlopainak páronkénti összehasonlítása 2D pontdiaagramok segítségével; formula argumentum is megengedett
coplot(y~x|a*b) 2D-s pontdiaagramok megjelenítése a és b értékei alapján
mosaicplot(x) mozaikábra; x gyakorisági táblázat grafikus megjelenítése
assocplot(x) Cohen-Friendly asszociációs ábra az x kétdimenziós gyakorisági táblázatra

Haladó adatkezelés

apply(X,MARGIN,FUN,...) a mátrix vagy adattábla X minden sorára (MARGIN=1) vagy oszlopára (MARGIN=2) meghívja a FUN(...) függvényt

lapply(X,FUN,...) a vektor vagy lista X minden elemére meghívja a FUN(...) függvényt; X hosszával egyező méretű listát ad

sapply(X,FUN,...) ua. mint fent, de vektort vagy mátrixot ad vissza

tapply(X,INDEX,FUN,...) az X vektor, az INDEX faktor vagy faktorok listája, ahol mindegyik faktor hossza X hosszával azonos; a faktorok minden különböző értékkombinációihoz tartozó X elemre meghívja a FUN(...) függvényt; ha teheti vektort vagy mátrixot ad vissza

by(X,INDICES,FUN,...) ua. mint tapply(), de X adattábla vagy mátrix, így minden oszlopra külön-külön kapunk egy (szébb) outputot

aggregate(x,by,FUN,...) ua. mint tapply(), de X vektor, adattábla vagy mátrix is lehet, valamint csak az x-ben előforduló by-beli faktorérték-kombinációkra végez összesítést

ifelse(test,yes,no) a test logikai feltételtől függően a yes vagy a no a visszatérési érték (vektorizált végrehajtással is), pl.: ifelse(1:5<3,1,2)

switch(EXPR,...) az első paraméter (numerikus vagy karakter skalár) alapján választ a többi arg-ból visszatérési értéket, pl.: switch("a",a=1,b=2)

merge(x,y,by) x és y adattáblák összefűlése a közös by oszlop azonos értékei alapján; eltérő oszlopnevek esetén by.x és by.y megadása szükséges; nem egyező sorok beillesztése a logikai a1l, a1l.x és a1l.y arg-kal

stack(x) a „széles” x adattáblából „hosszút” készít 2 oszloppal: value x num. adatai egymás alatt; ind faktor jelzi, melyik oszlopból származik az adat

unstack(x) a stack() inverze, x-ből „széles” adattáblát állít elő

reshape(data,varying,v.names,timevar,times,idvar,direction) a data adattábla „finomabb” átalakítása széles (direction="wide") vagy hosszú (direction="long") formára: ?reshape

Hagyományos grafika (Csomag: graphics)

▢ A rajzeszközök kezelése:

windows(width,height) grafikus ablak nyitása a képernyőn adott méretben
jpeg(filename,width=480,height=480) rajzolás állományba. Továbbá: bmp(), png() és tiff(), ill. pdf(), postscript(), továbbá: ?device

dev.list() a megnyitott grafikus eszközök listája: típus és sorszám

dev.cur() az aktuális grafikus eszköz, a rajzparancsok erre az eszközre írnak

dev.set(which) a which sorszámú grafikus eszközt teszi aktuálissá

dev.off(which=dev.cur()) a which sorszámú grafikus eszköz lezárása

graphics.off() minden grafikus eszköz bezárása

layout(mat,widths,heights) a rajzterület felosztása mátrix paraméterrel

layout.show() a rajzterület felosztásának megmutatása

Felosztáshoz még ld. mfrow grafikus paramétert, továbbá: ?screen

▢ Magas szintű rajzfüggvények (ld. még *Leíró statisztika 2. (grafika)* részt)

curve(expr,from,to,n) az expr kifejezés x függvénye; a [from,to] interv. felett n pontra expr grafikonját adja, pl. curve(x**2-1),curve(dt(x,9))

sunflowerplot(x,y) ua. mint a plot(), de a közel eső pontokat külön jelöli

pie(x,labels,radius=0.8,clockwise=F) kördiagram az x vektorból

fourfoldplot(x) az x 2x2-es mátrix; 2 dichotóm változó kapcsolatának rajza

interaction.plot(x.factor,trace.factor,response,fun=mean) a response-ra kétszemponτος interakció ábrázolása fun függvény alapján

plotMeans(response,factor1,factor2,error.bars) hasonló az interaction.plot()-hoz, de az átlagra rajzol. Csomag: Rcmdr

matplot(x,y,) kétváltozós pontdiagram az x és y mátrix azonos oszlopai alapján. Továbbá: az alacsony szintű matpoints() és matlines()

contour(x,y,z) kontúr rajzolása, kontúrvonalak hozzáadása

filled.contour(x,y,z) ua., de a területeket színekkel tölti ki; jelmagyarázat is
image(x,y,z) ua., x, y méretű kép létrehozása z-beli színek alapján

persp(x,y,z) felület rajzolása a 3D-s koordinátarendszerben

Továbbá: stars(), symbols(), termplot(), ts.plot() stb.

▢ Színek megadása színnevekkel, paletta sorszámmal és "#RRGGBB" alakban:

colors() összesen 657 szín angol neve, pl.: plot(x,col="red")

show.colors(type) színek és színnevek megjelenítése a type értékei ("singles", "shades" vagy "gray") alapján. Csomag: DAAG

palette(value) színpaletta beállítása és lekérdezése; a col argumentumban használható sorszám a paletta megfelelő színére vonatkozik

gray(level) szürke szín(ek) előállítása "#RRGGBB" alakban; level értéke 0 (fekete) és 1 (fehér) között lehet, pl.: barplot(x,col=gray(0:5/5))

rgb(red,green,blue,alpha,maxColorValue) RGB színkód előállítása, ha az alpha értékét megadjuk "#RRGGBBAA" -t ad, továbbá hcl(), hsv()

rainbow(n) a színskála n db színe "#RRGGBBAA" alakban; hasonló még a heat.colors(n), terrian.colors(n), topo.colors() és cm.colors()

Továbbá: col2rgb(), colorRamp() fv-ek, ill. RColorBrewer és dichromat csomagok

▢ A legtöbb rajzfüggvény közös argumentuma:

add=FALSE ha igazra állítjuk, akkor a létező ábrához hozzáadja a rajzot

axes=TRUE ha hamisra állítjuk, akkor nem rajzol tengelyeket és szegélyt

type="p" az ábra típusa, értékei: "p", "l", "b", "o", "h", "s", "S", "n"

xlim,ylim az x és y tengely ábrázolási tartományát határozhatjuk meg, pl.: xlim=c(-3, 3), ylim=c(0,10)

xlab,ylab az tengelyek címkéit határozhatjuk meg, pl.: xlab="eladási ár"

main,sub az ábra főcímét és alcímét határozhatjuk meg, pl.: main="15 ház"

Alacsony szintű grafikus parancsok

points(x,y,type),lines(x,y) pontok ill. vonalak rajzolása; x és y vektor

text(x,y,labels) a labels szöveg kiírása a rajzterület (x,y) pontjához

mtext(text,side=3,line=0,outer=F) szöveg elhelyezése a belső vagy külső (outer=T) margón a line szövegsorban a side oldalon (1=alsó)

segments(x0,y0,x1,y1) vonal rajzolása az (x0,y0) ponttól az (x1,y1) -ig

arrows(x0,y0,x1,y1,angle=30,code=2) ua. mint fent, de code alapján nyilat tesz a (x1,y1) ponthoz (2), (x0,y0)-hoz (1), mindkettőhöz (3)

abline(a,b),abline(h=y),abline(v=x),abline(reg) vonal rajzolása, ahol a az y tengely metszéspontja; b a meredekség; h és v a vízszintes ill. függőleges vonal koordinátája; reg regressziós objektum

rect(x1,y1,x2,y2) téglalap rajzolása a bal alsó és a jobb felső koordinátákkal

polygon(x,y) síkidom rajzolása a megadott koordináták alapján

legend(x,y,legend,ncol,fill,lty,lwd,pch) jelmagyarázat rajzolása az

(x,y) pozícióba, a legend a jelmagyarázat szövegét tartalmazza

title(main,sub,xlab,ylab,outer=F) szöveg hozzáadása az ábrához
box(which) terület szegélyezése, which értékei: "plot", "figure", "inner", "outer"

axis(side,at,label) a tengelyek megjelenését szabályozhatjuk, a side a tengely helyét jelenti: 1-lent, 2-balra, 3-fent, 4-jobbra

rug(x,ticksize=0.03,side=1) x vektor adatai kis függőleges vonalakban

▢ Interaktív grafikus függvények:

locator(n,type) pont vagy vonal elhelyezése a rajzterületen klikkeléssel, n a pontok maximális számát jelenti, type értéke "n", "p", "l", "o" lehet

identify(x,y,labels) a rajzterület adatpontjaihoz rendelhetünk címkét klikkeléssel; a pontok az x és y, a címkék a labels argumentumban

Grafikus paraméterek

par(...) grafikus paraméterek lekérdezése és beállítása, pl.: par("col") ill. par(col="skyblue")

▢ A grafikus paraméterek a következőek:

adj szöveg igazítása a text(), mtext() és a title() függvényekben. Értéke 0 és 1 közötti: 0- balra, 0.5-középre, 1-jobbra igazított.

bg a grafikus eszköz háttérszíne, pl.: bg="red" (az összes színnév: colors())

bty a rajzterület szegélyezése, lehetséges értékei: "o" (alapértelmezés), "l", "7", "c", "u", "j" és "n"

cex, cex.axis, cex.lab, cex.main, cex.sub rajzszimbólumok és szövegek méretét állíthatjuk be az alapértelmezetthez viszonyítva

col, col.axis, col.lab, col.main, col.sub rajzszimbólumok. vonalak és szövegek színét állíthatjuk be

font, font.axis, font.lab, font.main, font.sub a szövegek stílusa: 1- normál, 2-dőlt, 3-félkövér, 4-félkövér és dőlt

las a tengelycímkék írásirányát állíthatjuk be: 0-a tengelyekkel párhuzamos, 1- vízszintes, 2-tengelyekre merőleges, 3-függőleges

lty a rajzolandó vonal típusa; értéke 0 és 6 közötti egész, vagy a számoknak megfelelő sztring: 0-"blank", 1-"solid", 2-"dashed", 3-"dotted", 4-"dotdash", 5-"longdash", 6-"twodash";

lwd a rajzolandó vonal vastagsága, alapértelmezett értéke 1

mar, mai a belső margó beállítása szövegsorban ill. inch-ben mérve

mfrow, mfc0l egy c(sorokszáma, oszlopokszáma) alakú vektor, amelyek a grafikus eszközt a sorok és oszlopok száma alapján mátrixszerűen feldarabolja, a területeket pedig sor (mfrow) vagy oszlop folytonosan (mfc0l) számozza; a sorszámok sorrendjében jelennek meg a rajzok

oma, omi a külső margó beállítása szövegsorban ill. inch-ben mérve

pch a rajzszimbólum alakját határozza meg: 0-25-ig vehet fel értéket vagy tetszőleges karakter is lehet, pl.: par(pch=1) vagy par(pch="*")

ps a pontméret beállítása szövegek számára

tck, tc1 a tengelyek beosztásainak méretét szabályozhatjuk a rajzterület méretéhez (tck) vagy a szövegsor magasságához (tc1) viszonyítva

xaxt, yaxt az x ill. y tengely rajzolását tiltjuk, ha "n" az értéke

xlog, **ylog** logikai érték, amely ha igaz, akkor a megfelelő koordináta logaritmikuskálán mért

Trellis grafika (Csomag: lattice)

A **lattice** csomag rajzfüggvényei első argumentumnak az általános $y\sim x|f1*f2$ formulát várják. Szokásos argumentumok: **data** az adattábla neve, amelynek változóit a formulában használjuk; **subset** az adattábla sorainak szűrése logikai kifejezéssel; **panel1** saját panel függvény megadása az egyéni megjelenítéshez.

▢ A **lattice** csomag rajzfüggvényei:

histogram(~x) hisztogram rajzolása; ld. **hist()**

densityplot(~x) (empirikus) sűrűségfüggvény

qq(~x), **qqmath(~x)** QQ-ábra; ld. **qqplot()**

stripplot(y~x) 1D pontdiagram; ld. **stripchart()**

bwplot(y~x) dobozdiagram; ld. **boxplot()**

dotplot(y~x) kétdimenziós pontdiagram; ld. **dotchart()**

barchart(y~x) oszlopdiaagram; ld. **barplot()**

xypplot(y~x) kétdimenziós pontdiagram; ld. **plot()**

spiom(~x) kétdimenziós pontdiagramok mátrixa; ld. **pairs()**

parallel(~x) ld. **?parallel** példáit

contourplot(y~x) kontúr ábra; ld. **contour()**

levelplot(z~x*y|f1*f2) z-beli színek az x,y pontban; ld. **image()**

wireframe(z~x*y|f1*f2) 3D felületek; ld. **persp()**

cloud(z~x*y|f1*f2) 3D pontdiagram

Több rajzfüggvény kategorikus változóra: **?vcd**.

Programozás

függvénynév <- function(argumentum lista) {kifejezések} függvény létrehozása, pl.: **num.na<-function(x) { sum(is.na(x)) }**

return(érték) a függvény visszatérési értékét adja, ha hiányzik a függvényből, akkor az utolsó kiértékelés eredménye lesz a visszatérési érték (ld. fenti példa)

{kifejezések} blokk utasítás; egyetlen kifejezés esetén is használható

if(feltétel) {kifejezések} feltételes végrehajtás

pl.: **if(kor<18) { print("fiatal") }**

if(feltétel) {kif.1}else {kif.2} általános feltételes végrehajtás

pl.: **if(kor<18) { print("fiatal") } else { print("Nem fiatal") }**

for(változó in szekvencia) {kif.} ismételt végrehajtás

pl.: **for (i in 1:10) { print(i) }**

while(feltétel) {kif.} ismételt végrehajtás;

pl.: **i<-1; while(i<=10) { print(i); i<-i+1 }**

repeat {kif.} ismételt végrehajtás; a kilépéshez a **break** utasítás szükséges,

pl.: **x<-1; repeat { print(x); x<-x+1; if (x>10) break }**

break, **next** kilépés a ciklusból ill. a következő cikluslépés végrehajtása

do.call(függvénynév, arg.) alternatív függvényhívás

Statisztikai próbák

▢ Klasszikus tesztek folytonos változó(k)ra:

z.test(x,mu=0,stderr) egymintás u-próba. Csomag: **TeachingDemos**

t.test(x,mu=0) egymintás t-próba

t.test(x,y,var.equal), **t.test(y~f,var.equal)** kétmintás t-próba vagy Welch-próba a **var.equal** argumentum értékétől függően

t.test(x,y,paired=T) páros t-próba

cor.test(x,y,method) próba korreláció/asszociáció vizsgálatára; ld. **cor()**

oneway.test(formula,var.equal=F) egyszempontos varianciaanalízis

aov(formula) varianciaanalízis

lm(y~x) lineáris regresszióanalízis

lm(y~f) egyszempontos varianciaanalízis

lm(y~f1+f2) kétszempontos varianciaanalízis interakció nélkül

lm(y~f1*f2) kétszempontos varianciaanalízis interakcióval

lm(y~x+f) kovarianciaanalízis

lm(y~x1+x2+x3) többszörös lineáris regresszió

▢ ANOVA post hoc-tesztek

TukeyHSD(x) Tukey-próba az x modellobjektumra; minden párt összehasonlít; a **plot()** függvénnyel grafikusán is megjeleníthető eredményt ad

pairwise.t.test(x, g, p.adjust.method, paired) páronkénti összehasonlítások x vektor g -beli csoportjai között; a *p-érték* módosításának esetei: "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"; páros t-próbákat végez **paired=T**-val

SimTestDiff(data,grp,resp,type,base,ContrastMat) párhuzamos összehasonlítások; a type értékei: "Dunnnett", "Tukey", "Sequen", "Williams" stb. Csomag: **SimComp**

▢ Nemparaméteres tesztek:

wilcox.test(x, mu) Wilcoxon-próba; továbbá egzakt változat a **wilcox.test()** (coin csomag) és a **wilcox.exact()** (exactRankTests csomag)

wilcox.test(x,y) Mann-Whitney U-próba

wilcox.test(x,y,paired=T) páros Wilcoxon-próba

pairwise.wilcox.test(x,g,p.adjust.method,paired=F) páronkénti

Wilcoxon-próba a g-beli csoportok alapján a p-érték módosításával

kruskal.test(formula) Kruskal-Wallis-féle H-próba

friedman.test(formula) Friedman-próba; hívható mátrix argumentummal is

▢ Klasszikus tesztek diszkrét változó(k)ra:

binom.test(x,n,p=0.5) egzakt binomiális próba

mcnemar.test(x,y) McNemar-próba; x és y faktor, vagy az x mátrix és y=NULL

prop.test(x,n,p) próba populációbeli arányok vizsgálatára

fisher.test(x,y) Fisher-féle egzakt próba; x, y faktor,ha x mátrix,y=NULL

chisq.test(x,y) khi-négyzet próba

mh.test(formula) marginális homogenitás teszt. Csomag: **coin**

▢ Szórásokra vonatkozó tesztek:

var.test(x,y) F-próba, két minta szórásának egyezésére; x formula is lehet

levene.test(y,group) Levene-próba. Csomag: **car**

bartlett.test(x,...) Bartlett-próba, két vagy több minta szórásának egyezésére; formulát is használhatunk

▢ Eloszás alakjára vonatkozó tesztek:

ks.test(x,y) egy- vagy kétmintás Kolmogorov-Szmirnov-próba

shapiro.test(x) Shapiro-Wilk-próba a normalitás ellenőrzésére

mshapiro.test(U) többváltozós Shapiro-Wilk próba. Csomag: **mvnornctest**

▢ Néhány többváltozós statisztikai függvény:

prcomp() főkomponensanalízis

factanal() faktoranalízis

hclust(), **kmeans()** klaszteranalízis

lda(), **qda()** diszkriminanciaanalízis

nnet() neurális háló illesztése. Csomag: **nnet**

Statisztikai modellek vizsgálata

Az általános statisztikai modell formulája:

függő változó ~ független változó(k)

A független változók helyén használt operátorok speciális jelentéssel bírnak: pl.:

+ hatás hozzáadása; * a főhatások és az interakciók; -1 a konstans tag (intercept) eltávolítása a modellből

lm(formula) lineáris modell illesztése

aov(formula) varianciaanalízis

glm(model,correlation) illesztés általánosított legkisebb négyzetek módszerével. Csomag: **nlme**

nlm(formula,start) nemlineáris regresszió

gnlm(formula,start) nemlineáris regresszió általánosított legkisebb négyzetek módszerével. Csomag: **nlme**

glm(formula,family) általános lineáris modell illesztése

Továbbá: **gam()**, **lme()**, **lmer()**, **nlme()**, **loess()**, **tree()** stb.

▢ Általános függvények:

summary(object) itt: bővebb információ a modellről: az „eredmény”

plot(object) a modell 4 diagnosztikus ábráját jeleníti meg

anova(object,...) ANOVA táblázat megjelenítése; modellek összehasonlítása

update(object,formula) a modell formulájának a módosítása

coef(object) az illesztés együtthatóinak a meghatározása

fitted(object) a becslt értékeket tartalmazza

resid(object) a modell reziduumaikat tartalmazza

predict(object,...) predikció a modellel

confint(object) konfidencia intervallum a modell paramétereire

model.matrix(object) a modellmátrix kiírása

model.tables(x) a négyzetösszegek kiírása ANOVA esetén

AIC(object) Akaike kritérium és a Bayes kritérium meghatározása