



Przemysław Biecek

Przewodnik po pakiecie



Oficina Wydawnicza GiS

Argument `append` określa, czy wynik tej funkcji ma być dopisany do końca pliku (o ile plik istnieje), czy też ten wynik ma nadpisać ewentualną zawartość wskazanego pliku. Jeżeli wskazany plik nie istnieje, wynik jest w obu przypadkach taki sam. Do zapisu danych (wektora, macierzy lub ramki danych) w formacie tabelarycznym można wykorzystać funkcję `wri te . table()`. Poniższe polecenie zapisuje dane rozdzielając kolejne elementy znakiem tabulacji.

```
write.table(macierz, file="nazwa.pliku", sep="\t")
```

Duże i złożone obiekty lepiej przechowywać w postaci binarnej. Zapis w formacie plików binarnych umożliwia funkcja `save()`. Funkcja ta zapisuje wskazany obiekt lub listę obiektów w formacie `Rdata` (zazwyczaj pliki z rozszerzeniem `Rdata` lub `rda`). Do takiego formatu można zapisać nie tylko liczby, ale też złożone obiekty i funkcje. Jeżeli chcemy zapisać do pliku wartość wszystkich obiektów z przestrzeni nazw, to można skorzystać z funkcji `save . image()`. Zapisuje ona do pliku wszystkie dostępne obiekty. Podobny efekt ma polecenie z menu `File / Save workspace`.

1.7 Zadania do części „Łagodne wprowadzenie...”

W tym rozdziale przedstawiamy zbiór zadań do samodzielnego wykonania. Zadania zostały podzielone na trzy poziomy trudności, oznaczane liczbą liter `R`

`R` zadania podstawowe, bazujące wyłącznie na materiale z tej książki (dla większości użytkowników jest to zakres wystarczający do codziennej pracy z programem `R`),

`RR` zadania o podwyższonej trudności, dla biegłych użytkowników, wymagają często poszukania dodatkowych informacji w sieci lub w plikach pomocy `R`,

`RRR` zadania typu „sprawdź się”, najczęściej można je rozwiązać na wiele sposobów, rzecz w tym, by wybrać ten najlepszy.

Pod adresem <http://www.biecek.pl/R/odpowiedzi.R> znajdują się przykładowe rozwiązania poniższych zadań. Zapraszamy czytelników do zgłaszania własnych ciekawych rozwiązań i/lub zadań.

`R` Zadanie 1.1

Skonstruuj wektor kwadratów liczb od 1 do 100. Następnie używając operatora dzielenia modulo i funkcji `factor()` zlicz, które cyfry oraz jak często występują na pozycji jedności w wyznaczonych kwadratach.

`RR` Zadanie 1.2

Zbuduj własne tablice trygonometryczne. Przygotuj ramkę danych, w których zebrane będą informacje o wartościach funkcji sinus, cosinus, tangens i cotangens dla kątów: 0° , 30° , 45° , 60° , 90° . Zauważ, że funkcje trygonometryczne w `R` przyjmują argumenty w radianach.

R Zadanie 1.3

Przygotuj wektor 30 łańcuchów znaków następującej postaci: liczba.litera, gdzie liczba to kolejne liczby od 1 do 30 a litera to trzy duże litery A, B, C występujące cyklicznie.

R Zadanie 1.4

Wczytaj zbiór danych dane0 i napisz funkcję lub pętlę sprawdzającą typ i klasę każdej kolumny tego zbioru.

Patrz załącznik
z opisem
zbiorów danych.

R Zadanie 1.5

Z odczytanej ramki danych dane0 wyświetl tylko dane z wierszy o parzystych indeksach.

R Zadanie 1.6

Używając operatorów logicznych wyświetl ze zbioru danych tylko wiersze odpowiadające: pacjentkom starszym niż 50 lat u których wystąpiły przerzuty do węzłów chłonnych (cecha Wezly.chlonne=1).

R Zadanie 1.7

Wyświetl nazwy kolumn w zbiorze danych dane0, a następnie oblicz długość (liczbę znaków) nazw kolejnych kolumn.

R Zadanie 1.8

Napisz funkcję, która za argumenty przyjmie wektor liczb, a jako wynik zwróci trzy najmniejsze i trzy największe liczby. Jeżeli wejściowy wektor jest krótszy niż trzy liczby, to wyświetlany powinien być napis „za krótki argument”.

RR Zadanie 1.9

Zmodyfikuj funkcję z poprzedniego zadania tak, by otrzymywała też drugi argument i le, którym można określić liczbę skrajnych wartości wyznaczonych jako wynik. Domyślną wartością tego argumentu powinna być liczba 3.

RR Zadanie 1.10

Napisz funkcję poczatek() przyjmującą za pierwszy argument wektor, macierz lub ramkę a za drugi argument liczbę n. Niech to będzie przeciążona funkcja. Dla wektora powinna ona w wyniku zwracać n pierwszych elementów, dla macierzy i ramki danych powinna zwracać podmacierz o wymiarach $n \times n$.

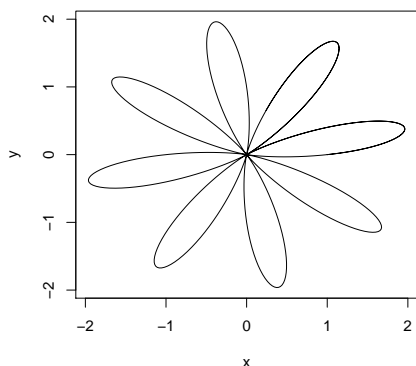
RR Zadanie 1.11 Narysuj funkcję, która w układzie biegunowym ma współrzędne:

$$r = 1 + \sin(t),$$

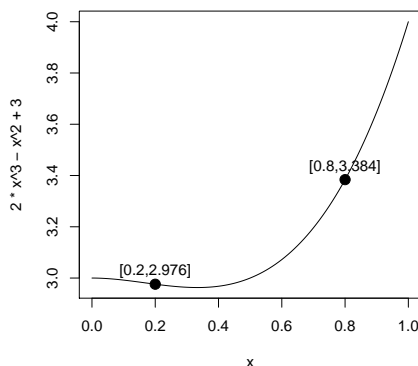
$$\phi = c * t.$$

dla $c = 1$, $c = 0.1$ i $c = 2.2$. Zobacz przykład na rysunku 1.6.

Wskazówka: trzeba zamienić współrzędne na układ kartezjański przekształceniami $x = r \cos(\phi)$ i $y = r \sin(\phi)$. Autor upierał się, że tego uczył w gimnazjum i takie podpowiedzi są zbędne, ale prawda jest taka, że autor nigdy nie był w gimnazjum. Przyp. żony.



RYSUNEK 1.6: Ilustracja do zadania 1.11



RYSUNEK 1.7: Ilustracja do zadania 1.12

R Zadanie 1.12

Używając instrukcji `curve()` narysuj wykres funkcji $f(x) = 2x^3 - x^2 + 3$ na przedziale $[0, 1]$.

Na wykresie tej funkcji zaznacz punkty odpowiadające wartościom w punktach 0.2 i 0.8, a następnie na wykres nanieś napisy opisujące współrzędne tych punktów. Zobacz przykład na rysunku 1.7.

R Zadanie 1.13

Pod adresem <http://www.biecek.pl/R/dane/daneBioTech.csv> znajduje się plik tekstowy z danymi. Dane są w formacie tabelarycznym, mają nagłówki, kolejne pola rozdzielane są średnikiem a kropką dziesiętną jest przecinek. Wczytaj te dane do programu R i przypisz je do zmiennej `daneBT`.

R Zadanie 1.14

Z odczytanych w poprzednim zadaniu danych wybierz tylko pierwsze trzy kolumny i pierwsze 10 wierszy. Zapisz ten fragment danych do pliku `maleDane.txt` na dysk `c:\` (użytkownicy Linuxa mogą zapisać do innego katalogu). Rozdzielaj kolejne pola znakiem tabulacji a kropką dziesiętną będzie kropka. Sprawdź w dowolnym edytorze tekstowym, co zapisało się do tego pliku.

RR Zadanie 1.15

Skonstruuj wektor 100 liczb, który jest symetryczny (tzn. elementy czytane od końca tworzą ten sam wektor co elementy czytane od początku). Pierwsze 20 liczb to kolejne liczby naturalne, następnie występuje 10 zer, następnie 20 kolejnych liczb parzystych (pozostałe elementy określone są przez warunek symetrii). Napisz funkcję, która sprawdza czy wektor jest symetryczny i sprawdź czy wygenerowany wektor jest symetryczny.

RRR Zadanie 1.16

Napisz funkcję `localMin()`, której argumentem będzie ramka danych, a wynikiem będą te wiersze, w których w jakiegokolwiek liczbowej kolumnie występuje wartość najmniejsza dla tej kolumny. Kolumny z wartościami nie-liczbowymi nie powinny być brane pod uwagę.

Innymi słowy jeżeli ramka ma trzy kolumny z wartościami liczbowymi, to wynikiem powinny być wiersze, w których w pierwszej kolumnie występują wartości minimalne dla pierwszej kolumny oraz wiersze, w których w drugiej kolumnie występują wartości minimalne dla drugiej kolumny oraz wiersze, w których w trzeciej kolumnie występują wartości minimalne dla trzeciej kolumny. Odczytaj ramkę danych z zadania 1.13 i sprawdź działanie napisanej funkcji.

R Zadanie 1.17

Poniższa funkcja nie działa poprawnie, powinna wyznaczać kwadraty kolejnych liczb ale tego nie robi. Skopiuj ją do programu R a następnie użyj instrukcji `fix()`, by poprawić funkcję `kwadratyLiczb()`.

```
kwadratyLiczb <- function(x) {
  1:x^2
}
```

R Zadanie 1.18

Funkcja `ecdf()`, wyznacza dystrybuantę empiryczną. Przyjrzyj się trzeciej linii z poniższego przykładu oraz spróbuj przewidzieć co jest wynikiem tego wyrażenia i jaka funkcja jest wywoływana jako druga.

```
data(iris)
x <- iris[,1]
ecdf(x)(x)
```

R Zadanie 1.19

Znajdź liczbę x z przedziału $[0 - 1]$ dla którego poniższe wyrażenie zwraca wartość `FALSE`.

```
x + 0.1 + 0.1 == 0.1 + 0.1 + x
```

R Zadanie 1.20

Dla zbioru danych `iris` narysuj wykres przedstawiający zależność pomiędzy dwoma wybranymi zmiennymi. Użyj funkcji `png()` i `pdf()` aby zapisać ten wykres do pliku.

TABELA 2.16: Inne przydatne funkcje systemowe z pakietu base.

<code>shell.exec()</code>	Uruchamia wskazany plik znajdujący się na lokalnym dysku lub wskazany przez adres URL, w programie skojarzonym z danym plikiem.
<code>shell()</code>	Uruchamia komendę w systemie operacyjnym. Działanie identyczne z poleceniem <code>system()</code> , choć to ponoć bardziej intuicyjny interfejs. Sposób podawania argumentów do obu funkcji wygląda praktycznie tak samo.
<code>browseEnv()</code>	Ta funkcja przygotowuje stronę HTML z opisem środowiska oraz przestrzeni nazw, w której pracujemy oraz otwiera tę stronę w domyślnej przeglądarce www.

2.10 Zadania do części „pazuRrry”

W tym rozdziale przedstawiamy zbiór zadań do samodzielnego wykonania. Zadania zostały podzielone na trzy poziomy trudności, oznaczane liczbą liter R. Pod adresem <http://www.biecek.pl/R/odpowiedzi.R> znajdują się przykładowe rozwiązania poniższych zadań. Zapraszamy czytelników do zgłaszania własnych ciekawych rozwiązań i/lub zadań.

R Zadanie 2.1

Odczytaj ramkę danych z zadania 1.13. Następnie zamień dane liczbowe z kolumny `Wiek` na zmienną czynnikową, dzieląc pacjentki na 3 grupy: o wieku do 45 lat, o wieku powyżej 55 lat i o wieku pośrednim. Poziomy tej zmiennej powinny nazywać się następująco: "wiek <45", "45<= wiek <=55", "wiek >55". Następnie wyświetl macierz kontyngencji dla tej zmiennej i dla pary zmiennych wieku oraz dla płci. Dodaj do macierzy sumy brzegowe. Wyświetl płaską macierz kontyngencji dla trójki zmiennych czynnikowych, dwóch powyższych i jeszcze zmiennej `WIT`.

RR Zadanie 2.2

Pod adresem <http://www.biecek.pl/R/dane/imieniny.txt> znajduje się plik tekstowy z imionami i datami imienin dla kolejnych imion. Plik jest w dosyć kłopotliwym formacie, mianowicie w każdym wierszu w pierwszej pozycji znajduje się imię, a po nim występują daty imienin. Wszystkie te pola rozdzielone są spacją. Ponieważ jednak różne imiona mają różne liczby imienin dane te nie są w postaci tabelarycznej. Odczytaj dane tak, by każdy wiersz był jednym elementem (można np. za separator wskazać `;` nie występuje on w tym pliku, cała linia zostanie więc traktowana jako jeden element). Sprawdź ile imion znajduje się w tym pliku z danymi.

RRR Zadanie 2.3

Po odczytaniu danych z zadania 2.2 użyj `strsplit()` i `sapply()`, by z odczytanych danych wydobyć tylko informacje o imionach. Zlicz liczby znaków w kolejnych imionach i wyznacz macierz kontyngencji opisującą ile imion ma określoną długość. Sprawdź, które imię ma najwięcej znaków i które imiona mają najmniej znaków. Większość imion żeńskich kończy się literą 'a',

Pamiętaj przy tym, że domyślnie pola tekstowe konwertowane są na zmienne czynnikowe.

wykorzystując tę informację zlicz liczbę żeńskich imion. Sprawdź ile liter rozpoczyna się literą A, ile B, ile C itp. Sprawdź, które imiona kończą się sufiksem *anna*. Polskie litery występujące w imionach zamień na odpowiedniki łacińskie, np. *ą* na *a*, *ź* na *z* itp. Sprawdź, ile imion zawierało polskie litery.

Od tej reguły są wyjątki np. żeńskie imię *Beatrycze* i męskie *Bonawentura*. Dla uproszczenia zapomnijmy o wyjątkach.

RR Zadanie 2.4

Po odczytaniu danych z powyższego zadania sprawdź, kto ma imieniny 30 października. Wyświetl te imiona w porządku leksykograficznym. Sprawdź, które imię ma najczęściej imieniny. Sprawdź, w którym dniu roku najwięcej imion ma imieniny. Sprawdź, w którym dniu miesiąca a następnie, w którym miesiącu najwięcej imion ma imieniny.

Na bazie wszystkich imion zrób analizę używalności poszczególnych liter. Która litera jest najpopularniejsza? Dlaczego?

R Zadanie 2.5

Odczytaj ramkę danych z zadania 1.13. Następnie używając funkcji `by()` wyświetl podsumowanie zmiennej `Wiek` osobno dla grupy `WIT=brak` i dla grupy `WIT=obecny`.

R Zadanie 2.6

Wyznacz wyznacznik, wartości własne oraz wektory własne macierzy:

$$\begin{bmatrix} 1 & 5 & 3 \\ 2 & 0 & 5 \\ 1 & 2 & 1 \end{bmatrix}.$$

RR Zadanie 2.7

Używając funkcji `outer()` zbuduj i wyświetl na ekranie tabliczkę mnożenia liczb od 1 do 10.

R Zadanie 2.8

Odczytaj ramkę danych z zadania 1.13. Następnie wyznacz histogram dla zmiennej `Wiek` i zapisz go do pliku `hist.pdf` w wymiarach 5×5 cali.

RRR Zadanie 2.9

Pod adresem `http://money.pl/` można odczytać aktualne wartości indeksów giełdowych. Wczytaj zawartość tej strony do programu R, a następnie wyciągnij z niej dane o nazwie i wartościach indeksów.

Zauważ, że w treści HTML tabela z nazwami i wartościami indeksów rozpoczyna się od linii `<table id="tgpw" class="tabela">`.

RR Zadanie 2.10

Używając profilera spróbuj przyspieszyć poniższy fragment kodu. Sprawdź, które instrukcje są czasochłonne i zastąp je szybszymi. Następnie sprawdź jak bardzo udało Ci się przyspieszyć ten kod.

Wskazówka: Funkcja `cor()` wywołuje masę dodatkowych funkcji, a tylko jedną niezbędną, czyli `Internal()`.

```
Rprof("profiler.out", interval = 0.01, memory.profiling=TRUE)
wyniki = NULL
for (i in 1:10000) {
  wektorX <- rnorm(20)
  wektorY <- rnorm(20)
  wyniki[i] <- cor(wektorX, wektorY)
}
Rprof(NULL)
summaryRprof("profiler.out", memory="both")
```

RR Zadanie 2.11

Napisz funkcję `test()`, której argumentem będzie komenda do sprawdzenia. Funkcja `test()` powinna wykonać daną komendę. Jeżeli zakończy się ona błędem to funkcja `test()` powinna wypisać na ekranie napis KLAPA, a jeżeli wykonanie będzie bezbłędne, to powinien pojawić się napis SUKCES. Przykładowe wywołanie `test(cor(1))` powinno spowodować wyświetlenie napisu KLAPA.

R Zadanie 2.12

Znajdź miejsca zerowe wielomianu $x^3 - 3x^2 - x + 3$. Znajdź najmniejszy wspólny dzielnik tego wielomianu i $x^3 - 12x^2 - x + 12$.

R Zadanie 2.13

Używając funkcji `head()` i `by()` dla zbioru danych z zadania 1.13 wyświetl trzy pierwsze wiersze danych dla mężczyzn i trzy pierwsze wiersze danych dla kobiet.

RR Zadanie 2.14

Jeżeli argumentem funkcji `diag()` jest macierz, to wynikiem jest wektor wartości z przekątnej. Napisz funkcję `diag2()` przyjmującą dwa argumenty. Pierwszym będzie macierz danych a drugim liczba. Wynikiem będzie przekątna przesunięta o wartość drugiego argumentu. Przykładowo, jeżeli drugi argument będzie miał wartość 0, to wynikiem będzie główna przekątna, jeżeli będzie miał wartość 1, to wynikiem będą elementy położone bezpośrednio nad główną przekątną, jeżeli wartość 2, to wyniki położone dwa wiersze ponad główną przekątną itp.

RR Zadanie 2.15

Napisz funkcję wyświetlającą nazwy 10 zmiennych zajmujących najwięcej pamięci w przestrzeni roboczej R.

R Zadanie 2.16

Pod adresem <http://www.biecek.pl/R/dane/daneBioTech.xls> znajduje się plik z danymi w formacie programu Excel. Skopiuj ten plik na dysk twardy, a następnie odczytaj dane z obu zakładek.

3.8 Zadania do części „Wybrane procedury statystyczne”

W tym rozdziale przedstawiamy zbiór zadań do samodzielnego wykonania. Zadania zostały podzielone na trzy poziomy trudności, oznaczane liczbą liter R. Pod adresem <http://www.biecek.pl/R/odpowiedzi> znajdują się przykładowe rozwiązania poniższych zadań. Zapraszamy czytelników do zgłaszania własnych ciekawych rozwiązań i/lub zadań.

R Zadanie 3.1

Napisz funkcję momenty(), która dla zadanego wektora liczb wyznaczy średnią, wariancję, skośność i kurtozę.

RR Zadanie 3.2

Ze zbioru danych dane0 wyciągnij zmienną VEGF i narysuj dla niej histogram (ile podziałów przyjąć?). Co o zmiennej mówi ten histogram? Zastosuj transformację, która ustabilizuje tę zmienną, a następnie narysuj histogram dla transformowanych danych. Na histogramie zamiast liczebności zaznacz prawdopodobieństwa, następnie dorysuj do tego wykresu ocenę gęstości wyznaczoną z użyciem funkcji density().

R Zadanie 3.3

W zbiorze danych dane0 znajduje się zmienna liczbowa Wiek i czynnikowa Receptory.estrogenowe. Narysuj wykres pudełkowy dla zmiennej Wiek w rozbiciu na zmienną Receptory.estrogenowe. Narysuj wykres skrzypcowy dla zmiennej Wiek w rozbiciu na poziomy zmiennej Niepowodzenia.

R Zadanie 3.4

Narysuj wykres rozrzutu dla zmiennych Wiek i VEGF (bez krzywej regresji nieparametrycznej). Czy wykres ten zmieni się, jeżeli zmienne będą podane w innej kolejności? Następnie zmienną VEGF narysuj w skali logarytmicznej. Narysuj też wykres rozrzutu dla tych zmiennych w rozbiciu na poziomy zmiennej Niepowodzenia.

R Zadanie 3.5

Wygeneruj 20 liczb z rozkładu wykładniczego i narysuj dla wygenerowanego wektora dystrybuantę empiryczną. Następnie zrób to samo dla wektora 200 i 2000 liczb. Używając funkcji MASS::fitdistr() oceń parametr rate dla wylosowanego wektora.

R Zadanie 3.6

Używając metod analizy wariancji sprawdź, średnia której ze zmiennych ilościowych w zbiorze danych dane0 różni się w zależności od poziomu zmiennej Niepowodzenia. Następnie zbuduj wektor zawierający p-wartości dla wyników z wykonanych analiz.

Zmienna Receptory.progesteronowe ma wartości na 4 poziomach. Wykonaj dla niej analizę wariancji oraz testy post hoc.

R Zadanie 3.7

Wykonaj model regresji liniowej dla zmiennych ciśnienie . skurczowe i wiek ze zbioru danych daneSoc.

R Zadanie 3.8

Odczytaj ramkę danych z zadania 1.13. Następnie wykonaj model regresji dla zmiennej Kreatynina.1, za zmienne zależne wybierz Kreatynina.7 i/lub Wiek.

Wyznacz dla tych zmiennych model regresji liniowej, oceń które zmienne są w istotnej statystycznie zależności, a następnie przedstaw graficznie zależność pomiędzy tymi zmiennymi.

R Zadanie 3.9

Wykorzystując wybrany test statystyczny sprawdź, która ze zmiennych ilościowych z ramki danych dane0 ma rozkład istotnie różny od normalnego.

RR Zadanie 3.10

Wyznacz moc wybranego testu normalności jako funkcję poziomu istotności dla alternatywy w postaci rozkładu jednostajnego i dla liczebności próby 50.

R Zadanie 3.11

Sprawdź, które zmienne ilościowe z ramki danych dane0 są ze sobą istotnie statystycznie skorelowane.

R Zadanie 3.12

W zbiorze danych mieszkania podziel zmienną cena na zmienną jakościową opisującą, czy cena jest wyższa od 175000, a zmienną powierzchnia podziel na zmienną jakościową opisującą, czy powierzchnia jest większa od $45m^2$.

Następnie wyznacz poznane współczynniki zgodności dla tych dwóch zmiennych jakościowych.

RR Zadanie 3.13

Napisz funkcję, która dla danego wektora liczb (zakładamy, że są to liczby dodatnie) sprawdza, czy po zastosowaniu transformacji logarytmicznej, odwrotnej lub pierwiastkowej dane nie są „bardziej normalne”. Jako wynik funkcja przekazuje wektor danych przekształconych albo nie (tak by uzyskać wektor najbardziej zbliżony do rozkładu normalnego) oraz informację, czy i jaką transformację zastosowano. Do badania odległości do rozkładu normalnego wykorzystać można odległość pomiędzy dystrybuantami. Uwaga: porównujemy z rodziną rozkładów normalnych a nie jednym rozkładem!

R Zadanie 3.14

W zbiorze danych dane0 zmienna Receptory.estrogenowe występuje na 4 poziomach, odpowiednio (-), (+), (++) , (+++). Wykonaj analizę warian-

cji dla następujących funkcji kontrastów: porównujący pacjentki bez tych receptorów (grupa (-)) z pozostałymi grupami, porównujący pacjentki z brakiem lub słabą ekspresją tych receptorów (grupy (-) i (+)) z pozostałymi grupami, porównujący pacjentki z bardzo wysoką ekspresją (grupa (+++)) z pozostałymi grupami. Uwaga: te kontrasty nie muszą być ortogonalne.

R Zadanie 3.15

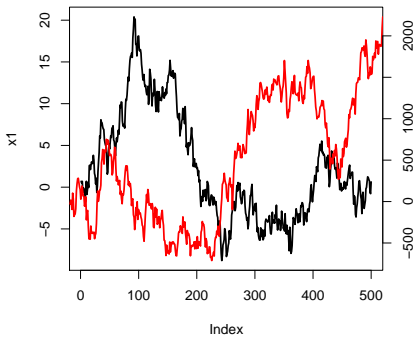
Sprawdź, czy w zbiorze danych dane0 zmienne Płeć i WIT wpływają na cechę Kreatynina. 7 w sposób addytywny, poprzez interakcję, czy też nie wpływają w żaden sposób. Zilustruj te zależności graficznie.

RRR Zadanie 3.16

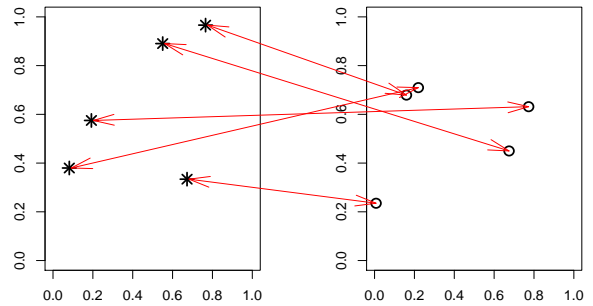
W celu testowania równości średnich w dwóch grupach zaimplementuj następującą dwustopniową procedurę. W pierwszym kroku sprawdź, czy w grupach zmienne mają rozkład normalny. Jeżeli tak, to porównaj zmienne testem t-Studenta, jeżeli nie, to porównaj testem Wilcoxona. Zbadaj na kilku przykładach, czy ta procedura kontroluje błąd pierwszego rodzaju oraz jaką ma moc w porównaniu do testów t-Studenta i Wilcoxona.

4.5 Zadania do części „gRrafika”

W tym rozdziale przedstawiamy zbiór zadań do samodzielnego wykonania. Zadania zostały podzielone na trzy poziomy trudności, oznaczane liczbą liter R. Pod adresem <http://www.biecek.pl/R/odpowiedzi> .R znajdują się przykładowe rozwiązania poniższych zadań. Zapraszamy czytelników do zgłaszania własnych ciekawych rozwiązań i/lub zadań.



RYSUNEK 4.98: Ilustracja do zadania 4.7



RYSUNEK 4.99: Ilustracja do zadania 4.8

R Zadanie 4.1

Dla pary zmiennych `cisnienie.skurczowe` i `cisnienie.rozkurczowe` ze zbioru danych `daneSoc` wyznacz histogram dwuwymiarowy. Aby poprawić czytelność wybierz odpowiednią liczbę klas.

R Zadanie 4.2

Dla pary zmiennych `Nowotwor` i `Wiek` ze zbioru danych `dane0` wyznacz wykres słonecznikowy.

R Zadanie 4.3

Narysuj macierz wykresów rozrzutu dla zbioru danych `mieszkania`.

RR Zadanie 4.4

Dla zmiennych `cisnienie.skurczowe` i `cisnienie.rozkurczowe` ze zbioru danych `daneSoc` wyznacz dwuwymiarową gęstość i narysuj wykres konturowy.

R Zadanie 4.5

Odczytaj ramkę danych z zadania 1.13. Następnie wyznacz macierz korelacji pomiędzy zmiennymi ilościowymi z tego zbioru danych. Przedstaw ją z użyciem wykresu `mapa ciepla` i wykresu `plotcorr()`

Użyj kilku poznanych funkcji. Który wykres jest najbardziej czytelny?

R Zadanie 4.6

W kolumnach 4, 5 i 6 w zbiorze danych z zadania 1.13 znajdują się poziomy zmiennej Kreatynina w trzech różnych chwilach czasowych. Używając funkcji `matplot()` narysuj zmiany tych zmiennych dla pierwszych 5 pacjentów.

RR Zadanie 4.7

Napisz funkcję, która za argumenty przyjmie dwa wektory liczb o równej długości. Następnie narysuje na jednym wykresie oba wektory (na osi OX będą kolejne liczby naturalne), każdy z własną osią OY (jedna oś narysowana powinna być po lewej stronie, druga po prawej). Osie, podobnie jak wykresy powinny być rysowane różnymi kolorami. Patrz przykład na rysunku 4.98.

RRR Zadanie 4.8

Narysuj dwa wykresy obok siebie. Na lewym wykresie narysuj pięć punktów o losowych współrzędnych (punkty zaznacz gwiazdkami), na prawym narysuj pięć punktów, również o losowych współrzędnych (zaznacz je okręgami). Następnie połącz strzałkami punkty z lewego wykresu z punktami z prawego wykresu (połączenie powinno być dokładne i niezależne od wymiarów urządzenia graficznego). Zobacz przykład na rysunku 4.99.

RR Zadanie 4.9

Ze strony <http://smarterpoland.pl/index.php/zbiory-danych/> pobierz dane dotyczące cen ofertowych sprzedaży używanych samochodów. Następnie przedstaw graficznie zależność pomiędzy wiekiem auta a ceną w podziale na różne marki / modele.

Wykonaj ten wykres używając najpierw funkcji z pakietu `graphics` a następnie `lattice` i `ggplot`. Z użyciem którego pakietu łatwiej było otrzymać wyniki o lepszej jakości?

RRR Zadanie 4.10

Ze strony <http://smarterpoland.pl/index.php/zbiory-danych/> pobierz dane dotyczące nagłówek RSS z różnych portali informacyjnych. Następnie przedstaw graficznie jak z czasem zmienia się częstość używania określonych słów w różnych portalach. Porównaj wyniki dla słowa `szok` i `polityk`.