

Technická univerzita v Košiciach  
Fakulta elektrotechniky a informatiky



## Aplikovaná štatistika

1. a 2. cvičenie - riešené príklady v R

Jana Petrillová, Mária Švecová

```

# PRÍKLAD 2.1

x=c(17, 18, 18, 22, 18, 17, 16, 17, 18, 18, 19, 18, 19, 17, 19, 19, 17, 18, 17, 18, 21, 17, 19, 18, 17, 20, 19,
    17, 18, 20, 19, 15, 19, 19, 19, 19, 18, 21, 18, 18)

# a)-----
vr= sort(x)                # variačný rad

# frekvenčná tabuľka (tabuľka rozdelenia početností)

xx=table(x)                # tabuľka absolútnych početností
ni=as.numeric(xx)         # absolutne početnosti z tabuľky
xi=unique(x)
xii=sort(xi)
n=length(x)
fi=ni/n
Ni=cumsum(ni)
Fi=cumsum(fi)
tab=cbind(xii,ni,Ni,fi,Fi)

# b)-----
# číselné charakteristiky

summ = summary(x)          # medián, prvý kvartil, tretí kvartil
# modus z frekvencnej tabuľky
ap = mean(x)               # aritmetický priemer
rozptyl = ((n-1)/n)*var(x) # rozptyl
s = sqrt(rozptyl)         # smerodajnú odchýlku alebo sqrt(rozptyl)
priem_odch = (1/n)*sum(abs(x-ap)) # priemernú odchýlku
var_koef = s/ap           # variačný koeficient
var_roz=max(x)-min(x)     # variačné rozpätie
med=as.numeric(summ[3])   # medián
prvy_kvartil=as.numeric(summ[2]) # prvý kvartil
treti_kvartil=as.numeric(summ[5]) # tretí kvartil
kvar_roz=treti_kvartil-prvy_kvartil # kvartilové rozpätie
koef_sik=(1/((s^3)*n))*sum((x-ap)^3) # koeficient šikmosti
koef_spic=(1/((s^4)*n))*sum((x-ap)^4)-3 # koeficient špicatosti
# alebo pomocou balicka e1071
koef_sik1=skewness(x)
koef_spic1=kurtosis(x)

info = data.frame(
  ap,
  rozptyl,
  s,
  priem_odch,
  var_koef,
  var_roz,
  med,
  prvý_kvartil,
  tretí_kvartil,
  kvar_roz,
  koef_sik,
  koef_spic
)
round(info,dig=3)          # vypise info a zaokruhli premenne info na tri desatinne miesta

# c)-----
# rôzne typy grafov

plot(xii, ni, col="blue", xlab="xi", pch=19) # bodový graf ni
plot(xii, ni, type="o", col="blue", xlab="xi", pch=19) # polygón ni
barplot(Ni, xlab="xi",col="blue", ylab="Ni") # stĺpcový graf Ni
hist(x, main = NULL, col = "blue", border = "white",xlab="xi") # histogram
pie(xii, labels = xii, col=gray(seq(from=0, to=1, length=5))) # kruhový graf
boxplot(x) # boxplot

# Bonus-----
# vytvori funkciu ktora najde modus
getmode = function(v) {
  uniqv = unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}
modus=getmode(x)

```

# PRÍKLAD 2.2

```
x=c(4.25, 4.26, 4.25, 4.07, 4.20, 4.40, 4.05, 4.00, 4.17, 3.96, 4.30, 4.22, 4.25, 4.15, 4.25, 4.25, 3.97, 4.10,
    4.47, 4.10, 4.17, 4.05, 4.25, 4.18, 4.05, 4.27, 4.03, 4.05, 4.15, 4.15, 4.40, 4.27, 4.25, 4.12, 4.30, 4.17,
    4.20, 4.17, 4.22, 4.32, 4.20, 4.05, 4.20, 4.37, 4.20, 4.25, 4.25, 4.15, 4.33, 4.10, 4.15, 4.25, 4.25, 4.05,
    4.32, 4.27, 4.12, 4.27, 4.50, 4.38, 4.27, 4.17, 4.17, 4.12, 4.25, 4.12, 4.30, 4.33, 4.33, 4.21, 4.12, 4.15,
    4.22, 4.20, 4.25, 4.50, 4.27, 4.10, 4.15, 4.31, 4.32, 4.10, 4.10, 4.17, 4.32, 4.02, 4.20, 4.17, 4.30, 4.44,
    4.17, 4.20, 4.45, 4.07, 4.27, 4.15, 4.12, 4.42, 4.21, 4.35)
```

# a)-----

# intervalové triedenie

```
min=min(x)
max=max(x)
n=length(x)
kk=round(sqrt(n)) # počet intervalov
hh=(max-min)/kk # šírka intervalov
k=11
h=0.05
```

# najdenie hraníc intervalov a ich absolútne početnosti

```
# bud
t=seq(3.95,4.5,by=0.05) # od 3,95 po 4,50 a skok je 0,05
zz=hist(x,seq(3.95,4.5,by=0.05), prob=FALSE)
ni=zz[["counts"]] # absolútne pocetnosti z hist
```

# alebo

```
histogram = hist(x,seq(3.95,4.5,by=0.05), plot=FALSE)
t = histogram$breaks # hranice intervalov z hist
ni = histogram$counts # absolútne pocetnosti z hist
cat(t,"\n",ni,"\n")
```

```
zi=(t[1:11]+t[2:12])/2 # triedny znak
```

# frekvenčná tabuľka (tabuľka rozdelenia početností)

```
fi=ni/n
Ni=cumsum(ni)
Fi=cumsum(fi)
tab=cbind(zi,ni,Ni,fi,Fi)
```

# b)-----

# číselné charakteristiky

```
ap = mean(x) # aritmetický priemer
rozptyl = ((n-1)/n)*var(x) # rozptyl
s = sqrt(rozptyl) # smerodajnú odchýlku alebo sqrt(rozptyl)
priem_odch = (1/n)*sum(abs(x-ap)) # priemernú odchýlku
var_koef = s/ap # variačný koeficient
var_roz=max(x)-min(x) # variačné rozpätie
med=4.15+h*((n+1)/2-33)/18 # medián
prvy_kvartil=4.10+h*((n/4)-19)/14 # prvý kvartil
tretí_kvartil=4.25+h*((3*n)/4-69)/12 # tretí kvartil
kvar_roz=tretí_kvartil-prvy_kvartil # kvartilové rozpätie
koef_sik=(1/((s^3)*n))*sum((x-ap)^3) # koeficient šikmosti
koef_spic=((1/((s^4)*n))*sum((x-ap)^4))-3 # koeficient špicatosti
# alebo pomocou balicka e1071
koef_sik=skewness(x)
koef_spic=kurtosis(x)
```

info = data.frame(

```
  ap,
  rozptyl,
  s,
  priem_odch,
  var_koef ,
  var_roz,
  med,
  prvý_kvartil,
  tretí_kvartil,
  kvar_roz,
  koef_sik,
  koef_spic
```

```
)
round(info,dig=3) # vypise info a zaokruhli premenne info na tri desatinne miesta
```

# modálny interval z frekvenčnej tabuľky

```
# modus
d11=ni[5]-ni[4]
d21=ni[5]-ni[6]
mod1=t[5]+h*(d11/(d11+d21))
d12=ni[6]-ni[5]
d22=ni[5]-ni[7]
mod2=t[6]+h*(d12/(d12+d22))
```

```
# c)-----  
# rôzne typy grafov  
  
plot(zi, ni, type="o", col="blue", xlab="zj",ylab="nj", pch=19) # polygón ni  
hist(x, main = NULL, col = "blue", border = "white",xlab="x") # histogram  
boxplot(x) # boxplot
```