

```

Lab1
#Oblicz
sin(2*pi)
cos(3/4)
tan(pi)
tan(sin(pi)/cos(pi))
log(100, 10)
log(15)
log(1/7, 7)
exp(3)
64^(1/3)

#wektory
wektor = seq(1,10)
x = seq(2,20,by=2)
length(x)
y = rev(x) #reverse
x*x
x^2
sqrt(sum(x^2)) #dlugosc euklidesowa
t(x)%*%y #mnozenie transpozycji przez wektor
x%*%t(y)

wektor2 = seq(5,10,length=13)
wektor2
z = c(1,2)
z1 = rep(c(1,2), times=5) #5ciokrotna replikacja z
z2 = rep(c(1,2), each=5) #5ciokrotna replikacja skladowych z
z1 + 4
z3 = z2[-c(length(z2))] #usuniecie ostatniej
c = z1 + z3
d = z1[z1>1]

#macierz
A = rbind(c(2,3,0),c(1,-1,2),c(1,1,-1))
A^2
A%*%A
t(A)
det(A)
solve(A)
b = A[3,]

x1 = c(1,3,7,12,14,23,33,41,48,50)
y1 = c
plot(x1,y1)
f = data.frame(x1,y1) #jakies polaczenie
plot(f)
c1 = rbind(x1,y1)
c2 = cbind(x1,y1)
plot(c1)
plot(c2)

curve(x^2+3*x-5, from = -3, to = 4 ) #funkcja kwadr w przedziale

```

Lab2

```
loty=read.csv("F:/Szkola/Statystyka/RStudio/loty.csv", sep=";")
oceny=read.csv("F:/Szkola/Statystyka/RStudio/oceny.csv", sep=";", dec=',')
truskawki=read.csv("F:/Szkola/Statystyka/RStudio/truskawki.csv", sep=";")

print(loty)
class(loty) #sprawdzenie typu
nazwy = names(loty)
#srednia
for(i in 1:6){
  print("Srednia")
  #Dla drugiej kolumny interpretacja: srednia liczba pasazerow w roku 1956 wynosila 328 osob
  print(nazwy[i])
  print(mean(loty[,i]))
}
#mediana
for(i in 1:6){
  print("Mediana")
  print(nazwy[i])
  print(quantile(loty[,i], probs=0.5))
  #Dla drugiej kolumny interpretacja: W 6 miesiacach liczba pasazerow pewnej lini lotniczej w roku 1956 byla mniejsza lub rowna
  315 osob i w 6 miesiacach liczba pasazerow pewnej lini lotniczej w roku 1956 byla wieksza lub rowna 315 osob"
}
#pierwszy kwantyl
for(i in 1:6){
  print("Pierwszy")
  print(nazwy[i])
  print(quantile(loty[,i], probs=0.25))
  #Dla drugiej kolumny interpretacja: W trzech miesiacach liczba pasazerow, w roku 1956 byla mniejsza lub rowna 301 osob i w
  9 miesiacach byla wieksza lub rowna 301 osob
}
#trzeci kwantyl
for(i in 1:6){
  print("Trzeci kwanty;")
  print(nazwy[i])
  print(quantile(loty[,i], probs=0.75))
}
for(i in 1:6){
  print("Odchylenie standardowe")
  print(nazwy[i])
  print(sd(loty[,i]))
  #Przecietnie liczba pasazerow odchyła sie od sredniej o 48 osób
}
for(i in 1:6){
  print("Zmiennosc")
  print(nazwy[i])
  z = sd(loty[,i])/mean(loty[,i])*100
  print(z)
  #Wystepuje slabe zroznicowanie liczby pasazerow w roku 1956
}

#HISTOGRAMY
min(loty)
max(loty)
```

```

przedzialy = seq(200,650,length=10)
par(mfrow=c(2,3))
kolory = c("red", "yellow", "pink", "blue", "green", "orange")
for(i in 1:6){
  hist(loty[,i],main=paste('Loty w', nazwy[i]), breaks=przedzialy, col=kolory[i], xlab='liczba pasazerow')
}

boxplot(loty[,1],loty[,2],loty[,3],loty[,4],loty[,5],loty[,6])

#srednia
apply(na.omit(oceny), 2,mean)
mean(na.omit(oceny[,2]))
print(oceny[,2])
#kwantyle
apply(na.omit(oceny), 2,quantile)
#odchylenie
apply(na.omit(oceny), 2,sd)
#zmiennosc
par(mfrow=c(2,2))
grupy = names(oceny)
for(i in 1:4){
  discrete.histogram(oceny[,i], freq=TRUE, main=grupy[i], xlabel=oceny)
}
boxplot(oceny[,1],oceny[,2],oceny[,3],oceny[,4])
par(mfrow=c(2,2))
for(i in 1:4){
  title=paste("wykres kolowy", grupy[i])
  pie(table(oceny[,i]))
}

print(truskawki)
print(mean(na.omit(truskawki[,2])))
print(mean(na.omit(truskawki[,1])))

print(quantile(na.omit(truskawki[,2])))
print(quantile(na.omit(truskawki[,1])))

print(sd(na.omit(truskawki[,2])))
print(sd(na.omit(truskawki[,1])))

#sporządź szeregi rozdzielcze przedziałowe plonów w poszczególnych latach (cut)
table(cut(truskawki[,1], 5))
table(cut(truskawki[,2], 5))

#przedstaw dane z szeregów rozdzielczych na wykresach kołowych;
par(mfrow=c(1,2))
for(i in 1:2){
  pie(table(cut(truskawki[,i], 5)))
}

przedzialy = seq(20,150, length=5)
for(i in 1:2){
  hist(truskawki[,i],breaks=przedzialy,freq=FALSE)
}

```

Lab3

#ZAD1 rozkład dwumianowy, Bin(n,p)

n = 5

x = 0:5

p = 0.3

prob = dbinom(x,n,p)

rbind(x,prob)

plot(x,prob,type="h", lwd = 50) #lwd to grubosc

dbinom(2,n,p) #gestosc -> wartosc dla podanego n

pbinom(2,n,p) #dystrybuanta -> $p(x=0) + p(x=1) + p(x=2)$

a = dbinom(3,n,p) #prawdopodobienstwo ze $X = 3 \rightarrow P(X=3)$

b = 1 - pbinom(2,n,p) # $P(X \geq 3) = P(X > 2)$

c = pbinom(2,n,p) # $P(X < 3) = P(X \leq 2)$

#ZAD2 rozkład dwumianowy, Bin(n,p)

n = 8

x = 0:8

p = 0.9

prob = dbinom(x,n,p)

rbind(x,prob)

a = dbinom(8,n,p)

b = dbinom(7,n,p)

c = 1 - pbinom(5,n,p)

d = sum(prob*x) #oczekiwana wartosc, jakiej mozemy sie spodziewac liczby zarowek lub $n*p$, przecietnie mozemy spodziewac sie ze 7 zarowek przekroczy zywtosc 500h

variance = sum(x^2*prob)-d^2

f = sqrt(variance) #odchylenie standardowe to $\sqrt{n*p*(1-p)}$

f2 = sqrt(n*p*(1-p)) #mozemy spodziewac sie ze przecietne odchylenie od sredniej wynosi jedna zarowka

#ZAD3 rozkład wykladniczy, typu ciaglego exp(lambda)

lambda = 0.01

curve(dexp(x,lambda), 0, 1000)

a = 1 - pexp(200,lambda) # $P(X \geq 200) = P(X > 200)$

b = pexp(100, lambda) # $P(X < 100) = P(X \leq 100)$

c = pexp(500, lambda)

#ZAD4 rozkład wykladniczy exp(lambda)

lambda = 1/2.4

curve(dexp(x,lambda), 0, 20)

a = 1 - pexp(3, lambda) # $P(X \geq 3) = P(X > 3)$

b = pexp(3,lambda) - pexp(2,lambda) # $P(a < X \leq b) = F(b) - F(a)$

f2 = function(x){x*lambda*exp(-lambda*x)}

integrate(f2,0,Inf)

#ZAD5 rozkład normalny, tez jest ciagly

mi = 0.13

sigma = 0.005

curve(dnorm(x,mi,sigma),mi-3*sigma,mi+3*sigma) #99,7% obserwacji na tym przedziale

pnorm(0.14,mi,sigma) - pnorm(0.12,mi,sigma)

#ZAD 6 I 7 DO DOMU BO CZEMU I NIE

mi = 120

sig = 15

curve(dnorm(x,mi,sig),mi-3*sig,mi+3*sig)

```
pnorm(135,mi,sig) - pnorm(111,mi,sig) #P(a < X <= b) = F(b) - F(a)
#ZAD7
mi = 46.8
sigma = 1.75
pnorm(50,mi,sigma) #P(X<=50)
1 - pnorm(48,mi,sigma) #P(X>=48)
#ZAD8 rozkład dwumianowy, gdy ponad 30 to duża próba s
p = 0.25
x = 0:100
n = 100
pbinom(15,n,p)
#przybliżenie rozkładem normalnym N(n*p, sqrt(n*p*q))
pnorm(15,n*p, sqrt(n*p*(1-p)))
```

```
#ZAD9 #Xsr ~ N(mi, sig/sqrt(n))
mi = 200
sigma = 10
n = 25
#P(199<xavr<202)
pnorm(202,mi,sigma/sqrt(25)) - pnorm(199,mi,sigma/sqrt(25))
#T = X1 + X2 + X3 + .... = N(n*mi, sqrt(n)* sigma)
pnorm(5100, n*mi, sqrt(n)*sigma)
```

```
#ZAD10
mi = 202
sigma = 14
n = 64
pnorm(206,mi,sigma/sqrt(n)) - pnorm(198,mi,sigma/sqrt(n))
#P(198<avr<206)
```

```
#ZAD11
mi = 0.5
sigma = 0.2
n = 100
1 - pnorm(47,mi*n,sigma*sqrt(n))
#suma zmiennych losowych
```

Lab4

#Przykład 5 z wykladuuuu

#P(phat<=232/100)

p = 0.25

n = 1000

T =232

proba = pnorm(T/n,p,sqrt(p*(1-p)/n))

dane = read.csv("F:/Szkola/Statystyka/RStudio/dane_est.csv", sep=";",dec=",")

print(dane)

print(dane[,1])

diamenty = na.omit(dane\$diamenty)

print(diamenty)

Populacja - wszystkie syntentyczne diamenty wyprudkowane nowa metoda

Proba - 12 syntentycznych diamentow wyprudkowanych nowa metoda

Badana zmienna - karaty

srednia = mean(diamenty)

wariancja = var(diamenty)

odchylenie = sd(diamenty)

PrzedzialUfnosciMU=function(srednia,odchylenie,sigma,liczebosc,ufnosc){

#srednia próby-X, odchylenie próby- S(a sigma to odchylenie dla populacji),

#liczebosc-n,

alfa=1-ufnosc

Lt=srednia-qt(1-alfa/2,liczebosc-1)*odchylenie/sqrt(liczebosc)

Pt=srednia+qt(1-alfa/2,liczebosc-1)*odchylenie/sqrt(liczebosc)

Lz=srednia-qnorm(1-alfa/2)*odchylenie/sqrt(liczebosc)

Pz=srednia+qnorm(1-alfa/2)*odchylenie/sqrt(liczebosc)

return(

if(liczebosc<30){

if(sigma==FALSE){print(paste("(",Lt,";",Pt,""))}}

else {print(paste("(",Lz,";",Pz,""))}}

}

else {print(paste("(",Lz,";",Pz,""))}}

)

}

PrzedzialUfnosciMU(srednia,odchylenie,FALSE,12,0.95)

t.test(diamenty, conf.level=0.95)

Z ufnoscia 0.95 przedzial (0.499;0.569) pokrywa nieznana srednia populacyjna mu

PrzedzialUfnosciSig2 = function(liczebosc, odchylenie, ufnosc){

alfa = 1 - ufnosc

if(liczebosc<30){

x1 = (liczebosc - 1)*odchylenie^2/qchisq(1-alfa/2,liczebosc-1)

x2 = (liczebosc - 1)*odchylenie^2/qchisq(alfa/2,liczebosc-1)

}

else{

x1 = (liczebosc - 1) + qnorm(1 - (1-ufnosc/2)) * sqrt(2*(liczebosc-1))

x2 = (liczebosc - 1) - qnorm(1 - (1-ufnosc/2)) * sqrt(2*(liczebosc-1))

}

return (print(paste(x1, x2)))

```

}
wynik = PrzedzialUfnosci dla Sig2(12, odchylenie, 0.95)

chi=sigma.test(diamenty,conf.level = 0.95)

#Z ufnoscia 0.95 przedzial (0.002;0.009) pokrywa nieznana prawdziwa wartosc warniacji dla populacji sigma^2

c=chi$conf.int
L=c[[1]]
P = c[[2]]
L_odchylenie = sqrt(L)
P_odchylenie = sqrt(P)
# Z ufnoscia 0.95 przedzial (0.039;0.094) pokrywa nieznana wartosc odchylenia standardowego dla populacji sigma

#ZAD2
mleko = na.omit(dane$mleko)
#populacja to wszystkie kobiety karmiace piersia
#proba to 20 kobiet karmiacych piersia
# badana zmienna to poziom pcb u kobiety
print(mleko)
srednia = mean(mleko)
wariancja = var(mleko)
odchylenie = sd(mleko)
t.test(mleko, conf.level=0.95)
PrzedzialUfnosciMU(srednia,odchylenie,FALSE,20,0.95)
#Z ufnoscia 0.95 przedzial (3.42;8.18) pokrywa nieznana rzeczywista srednia poziomu pcb w mleku wszystkich matek
chi = sigma.test(mleko,conf.level = 0.95)
#Z ufnoscia 0.95 przedzial (14.95;55.15) pokrywa nieznana rzeczywista wartosc warniacji zawartosci pcb w mleku wszystkich
matek karmiacych piersias
c=chi$conf.int
L=c[[1]]
P = c[[2]]
L_odchylenie = sqrt(L)
P_odchylenie = sqrt(P)
# Z ufnoscia 0.95 przedzial (3.86;7.43) pokrywa nieznana rzeczywista wartosc odchylenia standardowego zawartosci pcb w
mleku wszystkich matek karmiacych piersia

#ZAD3
papierosy = na.omit(dane$papierosy)
print(papierosy)
srednia = mean(papierosy)
wariancja = var(papierosy)
odchylenie = sd(papierosy)
z.test(papierosy,sigma.x = 0.7,conf.level=0.95) #WYLACZYC TEACHING EMOS. WLACZYC BDSA
PrzedzialUfnosciMU(srednia,0.7,TRUE,15,0.95)
#PODPUNKT B) blad estymacji - x_, blad estymacji + x_ => 2z * sig/sqrt(n) <= 0.3
z = qnorm(1-0.05/2)
szukane = (2*z * 0.7/0.3)^2
#Odchylenie standardowe zawiera sie w tym przedziale

#ZAD4
wodorosty = na.omit(dane$wodorosty)
srednia = mean(wodorosty)
wariancja = var(wodorosty)

```

```
odchylenie = sd(wodorosty)
t.test(wodorosty, conf.level=0.90)
chi = sigma.test(wodorosty, conf.level = 0.90) #teaching demos
```

```
#ZAD5
sygnał = na.omit(dane$sygnał)
odchlenieB=3
srednia=0
n=10
alfa=0.05
mean(sygnał)
z.test(sygnał, sigma.x=3, conf.level=0.95)
#ZAD6 rozkład nie jest podany
n = 1200
alfa = 0.05
alfa2 = alfa/2
srednia = 4.7 #mi
odchylenie = 2.2 #sigma
zsum.test(4.7, odchylenie, n, conf.level = 0.95)
#z ufnoscia 0.95 przedział (4.57;4.83) pokrywa nieznana rzeczywista srednia czasu wszystkich polaczen telefonicznych
#b) przedział ufnosci dla odchylenia standardowego
L = sqrt((n-1)*(odchylenie)^2/qchisq(0.95,n-1))
P = sqrt((n-1)*odchylenie^2/qchisq(alfa/2,n-1))
# z ufnoscia 0.95 przedział (2.11;2.3) pokrywa nieznana prawdziwa wartosc odchylenia standardowego dlugosci wszystkich
polaczen telefonicznych
```

```
#zad7
n = 365
srednia = 102
wariancja = 81
alfa = 0.02
zsum.test(102, sqrt(wariancja), n, conf.level = 0.98)
# z ufnoscia 0.98 przedział (100.9;103.1) pokrywa nieznana prawdziwa wartosc srednia zuzycia wody w fabryce
wariancja = 25
pewnosc = 0.95
L = sqrt((n-1)*81/qchisq(0.98,n-1))
P = sqrt((n-1)*81/qchisq(alfa/2,n-1))
```

```
#ZAD8 rozkład normalny
alfa = 0.05
sigma = sqrt(25)
#n=?
kw=qnorm(1-alfa/2)
(kw*sigma/1)^2
```

```
#ZAD9 rozkład normalny
odchylenie = 0.3
alfa = 0.1
kw=qnorm(1-alfa/2)
(kw*odchylenie/0.1)^2
alfa2 = 0.01
kw2=qnorm(1-alfa2/2)
(kw2*odchylenie/0.1)^2
#ZADd10
```



```
p = 0.015
phat = 4/100
n = 100
ufnosc = 0.95
L = phat - qnorm(1-alfa/2)*sqrt(phat*(1-phat))/sqrt(n)
P = phat + qnorm(1-alfa/2)*sqrt(phat*(1-phat))/sqrt(n)
prop.test(4,100,0.95)
#pokrywa nieznana rzeczywista proporcje wszystkich niedopelnionych puszek
```

```
#zad11
n = 120
blad = 24
alfa = 0.1
phat = 24/120
L = phat - qnorm(1-alfa/2)*sqrt(phat*(1-phat))/sqrt(n)
P = phat + qnorm(1-alfa/2)*sqrt(phat*(1-phat))/sqrt(n)
```

```
#zad12
alfa = 0.02
#a
p = 0.3
#blad oszacowania
#z * sqrt(phat(1-phat)/sqrt(n))
szukana = (qnorm(1-alfa/2) * sqrt(p*(1-p)))/0.05^2
```

```
#b
p = 0.5
szukana = (qnorm(1-alfa/2) * sqrt(p*(1-p)))/0.05^2
```