

**KAUNO TECHNOLOGIJOS UNIVERSITETAS**

**(Informatikos) FAKULTETAS**

**Vytenis Kriščiūnas, IFF-1/1**

**72 variantas**

**P160B003 Tikimybių teorijos ir statistikos**

**2-ojo individualaus laboratorinio darbo ataskaita**

Kaunas, 2022

# UŽDUOTIS

## R PROGRAMOS KODAS

attach(duomenys)

x = duomenys$variantas\_72\_x

y = duomenys$variantas\_72\_y

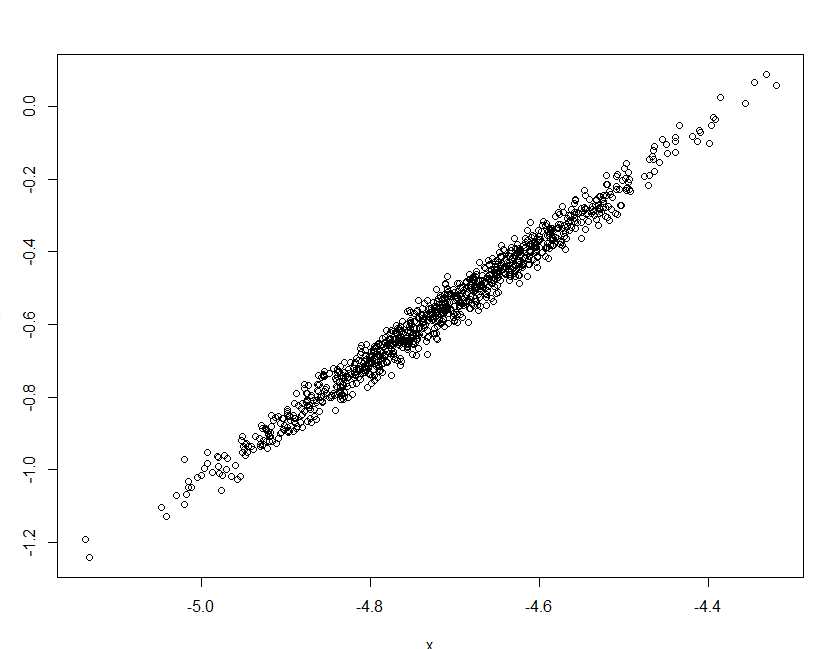
#1.1

#Kintamieji x ir y yra stipriai priklausomi, jų taškai sudaro vis augančią tiesę

plot(x,y)

## REZULTATAI IR IŠVADOS

Kintamieji x ir y yra stipriai priklausomi, jų taškai sudaro vis augančią tiesę.



1 Pav. Kintamųjų x-y grafikas

## R PROGRAMOS KODAS

#1.2

#Labai stipri teigiama tiesinė koreliacija

cor(x,y, method = 'pearson')

## REZULTATAI IR IŠVADOS

> cor(x,y, method = 'pearson')

[**1**] **0.9898373**

Labai stipri teigiama tiesinė koreliacija.

## R PROGRAMOS KODAS

#1.3

#p-value reikšmė yra daug mažesnė už standartinį reikšmingumo lygmenį 0.05

#Nulinę hipotezę, kad Pirsono koreliacijos koficiantas lygus nuliui reikėtų atmesti ir pereiti prie alternatyviosios hipotezės

cor.test(x, y)

## Rezultatai ir išvados

> cor.test(x, y)

Pearson's product-moment correlation

data: x **and** y

t = **219.9**, df = **998**, p-value < **2.2e-16**

alternative hypothesis: true correlation **is** **not** equal to **0**

**95** percent confidence interval:

**0.9885017** **0.9910184**

sample estimates:

cor

**0.9898373**

p-value reikšmė yra daug mažesnė už standartinį reikšmingumo lygmenį 0.05.

Nulinę hipotezę, kad Pirsono koreliacijos koficiantas lygus nuliui reikėtų atmesti ir pereiti prie alternatyviosios hipotezės.

## R PROGRAMOS KODAS

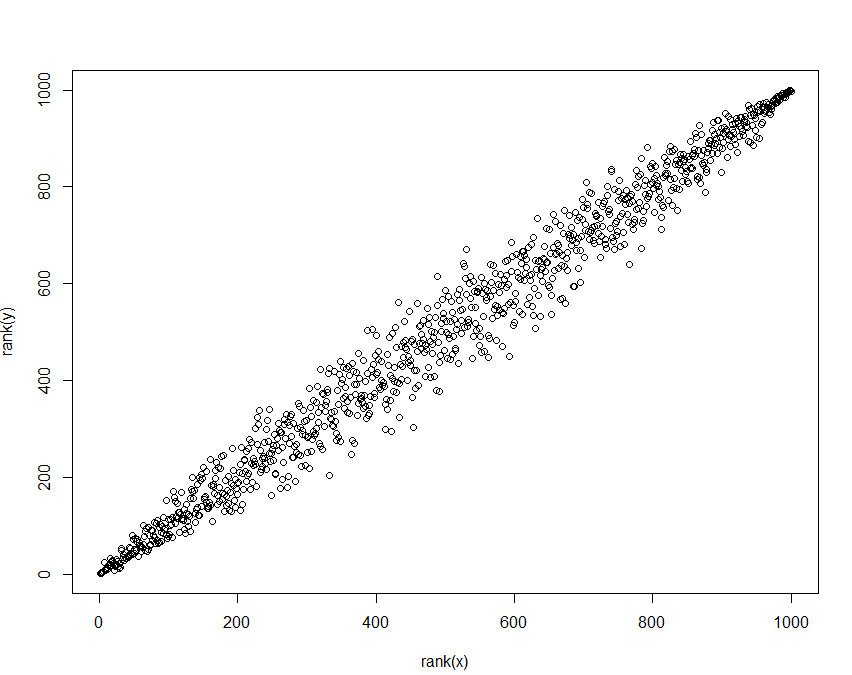
#1.4

#Priklausomybė tarp rangų yra ganėtinai didelė, tačiau taškai rangų grafe yra labiau išsisklaidę nei kintamųjų grafe

plot(rank(x), rank(y))

## Rezultatai ir išvados

Priklausomybė tarp rangų yra ganėtinai didelė, tačiau taškai rangų grafe yra labiau išsisklaidę nei kintamųjų grafe.



2 Pav. Kintamųjų x-y rangų grafikas

## R PROGRAMOS KODAS

#1.5

#Labai stipri teigiama tiesinė koreliacija

cor(x, y, method = 'spearman')

## Rezultatai ir išvados

> cor(x, y, method = 'spearman')

[**1**] **0.9890307**

Labai stipri teigiama tiesinė koreliacija.

## R PROGRAMOS KODAS

#1.6

#p-value reikšmė yra daug mažesnė už standartinį reikšmingumo lygmenį 0.05

#Nulinę hipotezę, kad Spirmeno koreliacijos koficiantas lygus nuliui reikėtų atmesti ir pereiti prie alternatyviosios hipotezės

cor.test(x, y, method = 'spearman')

## Rezultatai ir išvados

> cor.test(x, y, method = 'spearman')

Spearman's rank correlation rho

data: x **and** y

S = **1828218**, p-value < **2.2e-16**

alternative hypothesis: true rho **is** **not** equal to **0**

sample estimates:

rho

**0.9890307**

p-value reikšmė yra daug mažesnė už standartinį reikšmingumo lygmenį 0.05.

Nulinę hipotezę, kad Spirmeno koreliacijos koficiantas lygus nuliui reikėtų atmesti ir pereiti prie alternatyviosios hipotezės.

# UŽDUOTIS.

## R PROGRAMOS KODAS

x = duomenys$variantas\_72\_x

y = duomenys$variantas\_72\_y

#2.1

# k = 1.622, b = 7.085, lygtis y = 1.622 \* x + 7.085

plot(x,y)

lm(y~x)

k = **1.622**

b = **7.085**

lines(x, k\*x +b)

## Rezultatai ir išvados

> plot(x,y)

> lm(y~x)

Call:

lm(formula = y ~ x)

Coefficients:

(Intercept) x

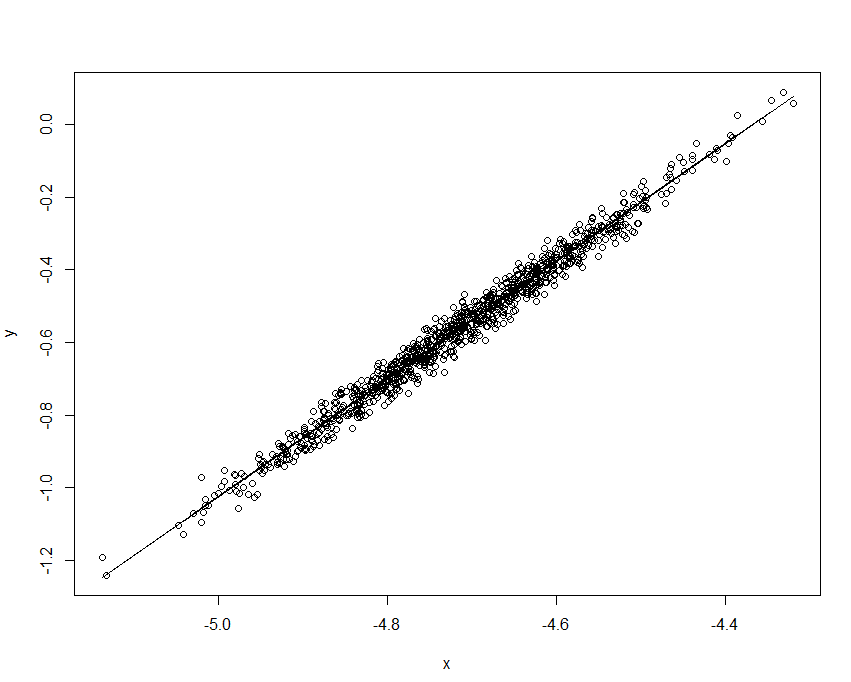
**7.085** **1.622**

> k = **1.622**

> b = **7.085**

> lines(x, k\*x +b)

k = 1.622, b = 7.085, lygtis y = 1.622 \* x + 7.085.



3 Pav. Numbrėžta regresijos lygties tiesė kintamųjų x-y grafike

## R PROGRAMOS KODAS

#2.2

#Tikrinsiu reikšmę 10

#Gautas atsakymas iš tiesės lygties: 23.305 yra labai panašus į spėjamą atsakymą: 23.30998

yReg = k \* **10** + b

yReg

predict(lm(y~x), data.frame(x=**10**))

## Rezultatai ir išvados

> yReg = k \* **10** + b

> yReg

[**1**] **23.305**

> predict(lm(y~x), data.frame(x=**10**))

**1**

**23.30998**

Tikrinsiu reikšmę 10.

Gautas atsakymas iš tiesės lygties: 23.305 yra labai panašus į spėjamą atsakymą: 23.30998.

## R PROGRAMOS KODAS

#2.3

#Nurodyta p-value reikšmė yra daug mažesnė už duotą reikšmingumo lygmenį 0.05

#Nulinę hipotezę, kad tiesinės lygties krypties koeficientas lygus nuliui atmentame, todėl galime pereiti prie alternatyviosios hipotezės

summary(lm(y~x))

## Rezultatai ir išvados

> summary(lm(y~x))

Call:

lm(formula = y ~ x)

Residuals:

Min **1**Q Median **3**Q Max

-**0.088418** -**0.020003** **0.000416** **0.020053** **0.086166**

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) **7.085498** **0.034815** **203.5** <**2e-16** \*\*\*

x **1.622448** **0.007378** **219.9** <**2e-16** \*\*\*

---

Signif. codes: **0** ‘\*\*\*’ **0.001** ‘\*\*’ **0.01** ‘\*’ **0.05** ‘.’ **0.1** ‘ ’ **1**

Residual standard error: **0.03002** on **998** degrees of freedom

Multiple R-squared: **0.9798**, Adjusted R-squared: **0.9798**

F-statistic: **4.835e+04** on **1** **and** **998** DF, p-value: < **2.2e-16**

Nurodyta p-value reikšmė yra daug mažesnė už duotą reikšmingumo lygmenį 0.05.

Nulinę hipotezę, kad tiesinės lygties krypties koeficientas lygus nuliui atmentame, todėl galime pereiti prie alternatyviosios hipotezės.

## R PROGRAMOS KODAS

#2.4

#Tiesinė regresijos determinacijos koeficinetas lygus 0.9798

#Šis koeficientas yra labai arti 1, todėl tiesinės regresijos modelis yra tinkamas šiems duomenims

summary(lm(y~x))

## Rezultatai ir išvados

> summary(lm(y~x))

Call:

lm(formula = y ~ x)

Residuals:

Min **1**Q Median **3**Q Max

-**0.088418** -**0.020003** **0.000416** **0.020053** **0.086166**

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) **7.085498** **0.034815** **203.5** <**2e-16** \*\*\*

x **1.622448** **0.007378** **219.9** <**2e-16** \*\*\*

---

Signif. codes: **0** ‘\*\*\*’ **0.001** ‘\*\*’ **0.01** ‘\*’ **0.05** ‘.’ **0.1** ‘ ’ **1**

Residual standard error: **0.03002** on **998** degrees of freedom

Multiple R-squared: **0.9798**, Adjusted R-squared: **0.9798**

F-statistic: **4.835e+04** on **1** **and** **998** DF, p-value: < **2.2e-16**

Tiesinė regresijos determinacijos koeficinetas lygus 0.9798.

Šis koeficientas yra labai arti 1, todėl tiesinės regresijos modelis yra tinkamas šiems duomenims.

## R PROGRAMOS KODAS

#2.5

#Iš histogramos galima matyti, kad liekanų skirstinys primena normalųjį skirstinį

#Iš grafiko galima matyti, kad skirstinys yra homoskedatiškas.

e = y - k\*x - b

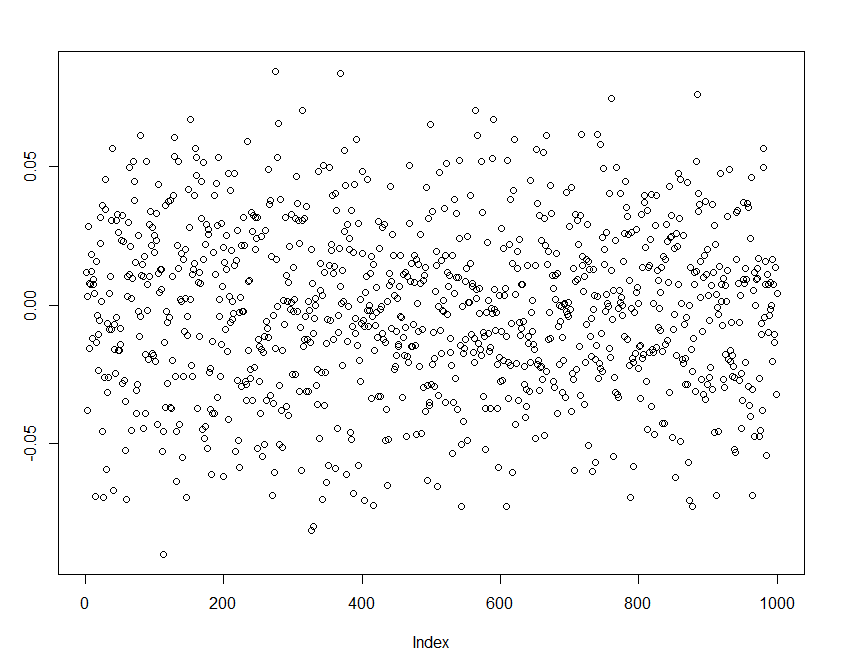
plot(e)

hist(e)

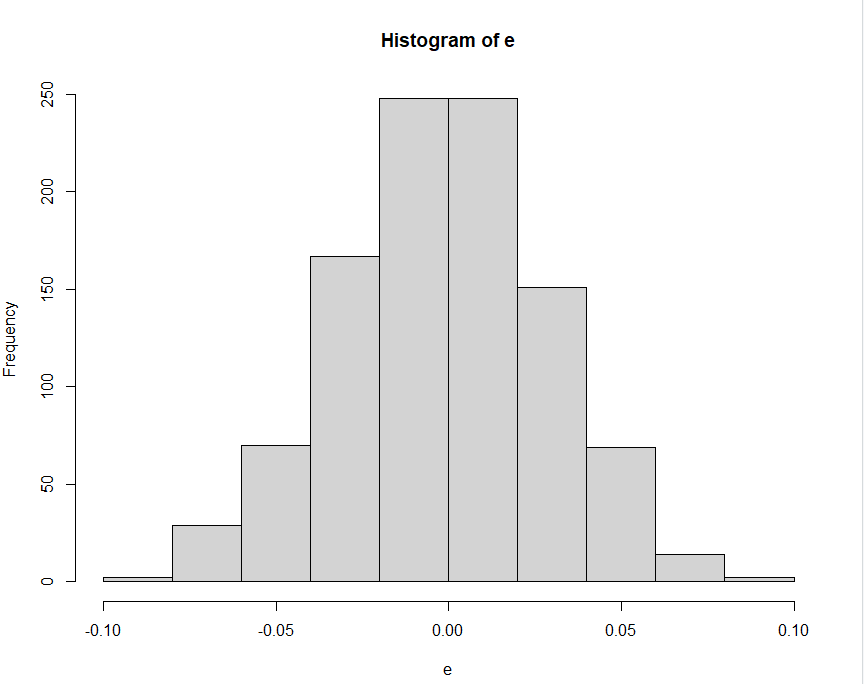
## Rezultatai ir išvados

Iš histogramos galima matyti, kad liekanų skirstinys primena normalųjį skirstinį

Iš grafiko galima matyti, kad skirstinys yra homoskedatiškas.



4 Pav. Liekanų grafikas



5 Pav. Liekanų histograma

## R PROGRAMOS KODAS

#2.6

#P-value = 0.4009 reikšmė yra didesnė už 𝛼 = 0.01, todėl neatmetame statistinės hipotezės, jog regresijos liekanų skirstinys yra normalusis.

library(MASS)

fitdistr(e, 'normal')

ks.test(e, 'pnorm', **0**, **0.0299946010** )

## Rezultatai ir išvados

> library(MASS)

> fitdistr(e, 'normal')

mean sd

-**0.0016171823** **0.0299946010**

( **0.0009485126**) ( **0.0006706997**)

> ks.test(e, 'pnorm', **0**, **0.0299946010** )

Asymptotic one-sample Kolmogorov-Smirnov test

data: e

D = **0.028276**, p-value = **0.4009**

alternative hypothesis: two-sided

P-value = 0.4009 reikšmė yra didesnė už 𝛼 = 0.01, todėl neatmetame statistinės hipotezės, jog regresijos liekanų skirstinys yra normalusis.