



**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**INFORMATIKOS FAKULTETAS**

**LUKAS KUZMICKAS**

Studijų modulis

**P160B003 Tikimybių teorija ir statistika**

**1 laboratorinio darbo ataskaita**

Kaunas, 2022

## DARBO TIKSLAS

Ugdyti gebėjimus taikyti teorines statistikos žinias praktikoje, programuoti statistikos uždavinius, panaudojant R programavimo kalbą, atlikti tiriamąją duomenų analizę, tikrinti hipotezes, apskaičiuoti parametrų pasikliautuosius intervalus, interpretuoti gautus rezultatus, formuluoti išvadas, rengti ataskaitas.

## 1. UŽDUOTIS.

### 1. APRAŠOMOJI STATISTIKA IR NEPARAMETRINĖS HIPOTEZĖS

Iš duomenų imties „duomenys\_hipoteziu\_tikrinimui.csv“ nuskaitykite savo varianto duomenis tokiu būdu:

```
N=variantas#Reikia įrašyti savo varianto numerį N
data=read.csv('duomenys_hipoteziu_tikrinimui.csv',header=TRUE) #Failas turi būti darbiname kataloge
data=data[1:1000,N] #Duomenys, kuriuos reikia analizuoti – imtis.
```

**1.1.** Raskite skaitines imties charakteristikas: imties dydį, minimumą, maksimumą, imties plotį, kvartilius, kvartilių skirtumą (IQR), empirinį vidurkį, pataisytą dispersiją, pataisytąjį standartinį nuokrypį.

**1.2.** Nubraižykite histogramą ir stačiakampę diagramą. Pagal gautus grafikus, parinkite vieną iš trijų tolydžių skirstinių (normalųjį, eksponentinį arba tolygųjį), kuris galėtų geriausiai tikti duotiesiems duomenims.

**1.3.** Raskite skirstinio parametrų taškinis įverčius.

**1.4.** Suformuluokite ir patikrinkite atitinkamą suderinamumo hipotezę (reikšmingumo lygmenį  $\alpha$  pasirinkite patys). Užrašykite nagrinėjamo populiacijos požymio skirstinio funkciją.

1 pav. Pirmosios užduoties sąlyga

## 1.1. R PROGRAMOS KODAS

```
#1 UŽDUOTIS
N=40 #variantų numeris
data=read.csv('duomenys_hipoteziu_tikrinimui.csv', header=TRUE)
data = data[1:1000,N]
data
#1.1. Raskite skaitines imties charakteristikas: imties dydį,
#minimumą, maksimumą, imties plotį, kvartilius,
#kvartilių skirtumą (IQR), empirinį vidurkį,
#pataisytą dispersiją, pataisytą standartinį nuokrypį.

n = length(data)
#imties dydis
n
data_minimum = min(data)
#Minimumas
data_minimum
data_maximum = max(data)
#Maximumas
data_maximum
#imties plotis
data_plot = diff(range(data))
data_plot
#kvartiliai
quantile(data, c(0.25, 0.5, 0.75), type = 2)
#kvartilių skirtumas
KP <- IQR(data)
KP
#Empirinis vidurkis
emp_vidurkis = mean(data)
emp_vidurkis
#Pataisytą dispersiją
disp_pat <- var(data)
disp_pat
#Pataisytą standartinį nuokrypį
standartas_pat <- sd(data)
standartas_pat
```

2 pav. 1.1. užduoties kodas

## REZULTATAI IR IŠVADOS

```
> #1.1. Raskite skaitines imties charakteristikas: imties dydį,  
> #minimumą, maksimumą, imties plotį, kvartilius,  
> #kvartilių skirtumą (IQR), empirinį vidurkį,  
> #pataisytąją dispersiją, pataisytąjį standartinį nuokrypį.  
>  
> n = length(data)  
> #Imties dydis  
> n  
[1] 1000  
> data_minimum = min(data)  
> #Minimumas  
> data_minimum  
[1] -1.820401  
> data_maximum = max(data)  
> #Maximumas  
> data_maximum  
[1] 7.022358  
> #imties plotis  
> data_plot = diff(range(data))  
> data_plot  
[1] 8.842759  
> #kvartiliai  
> quantile(data, c(0.25, 0.5, 0.75), type = 2)  
      25%      50%      75%  
0.4617784 2.5787572 4.7671800  
> #kvartiliu skirtumas  
> KP <- IQR(data)  
> KP  
[1] 4.303099  
> #Empirinis vidurkis  
> emp_vidurkis = mean(data)  
> emp_vidurkis  
[1] 2.594081  
> #Pataisytoji dispersija  
> disp_pat <- var(data)  
> disp_pat  
[1] 6.356343  
> #Pataisytasis standartinis nuokrypis  
> standartas_pat <- sd(data)  
> standartas_pat  
[1] 2.521179
```

3 pav. 1.1. užduoties rezultatai

Rezultatai:

- Imties dydis – 1000 elementų.;
- Imties minimumas = -1.820401;
- Imties maksimumas = 7.022358;
- Imties plotis = 8.842759;
- Kvantiliai (25%,50%,75%) – kiek reikšmių didesnių;
- Kvantiliai: 25% (0.4617784);
- Kvantiliai: 50% (2.5787572);
- Kvantiliai: 75% (4.7671800);
- Kvartilių skirtumas (IQR) = 4.303099;
- Empirinis vidurkis = 2.594081;
- Pataisytoji dispersija = 6.356343;
- Pataisytasis standartinis nuokrypis = 2.521179.

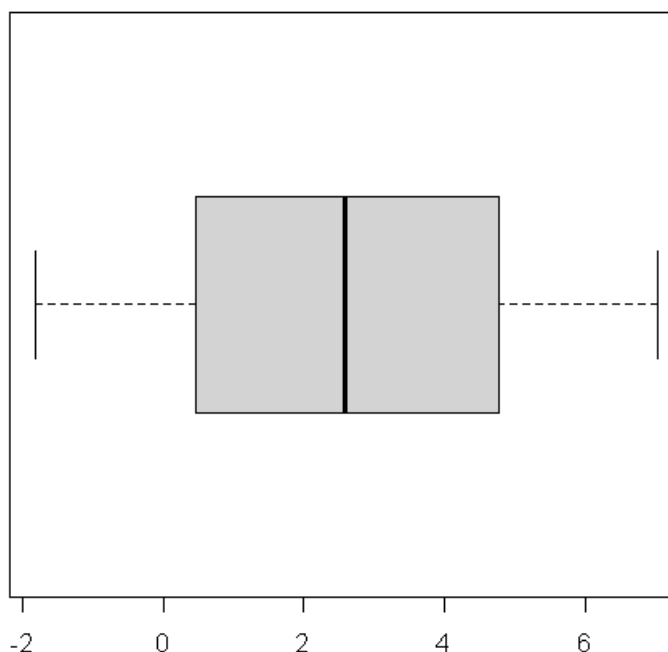
## 1.2.R PROGRAMOS KODAS

```
#1.2. Nubraižykite histogramą ir stačiakampę diagramą.  
#Pagal gautus grafikus, parinkite vieną iš trijų tolydžių  
#skirstinių (normalųjį, eksponentinį arba tolygųjį),  
#kuris galėtų geriausiai tiktį duotiesiems duomenims.  
  
#Stačiakampė diagrama  
boxplot(data, horizontal = TRUE, main = "Stačiakampė diagrama")  
k <- ceiling(log10(length(data))*3.32+1)  
k  
  
#Histograma  
hist(data, breaks = seq(min(data), max(data), length = k+1),  
      probability = T,  
      right = F,  
      col = grey(0.9),  
      ylab = "Santykiniai dažniai",  
      main = "Histograma")  
)  
#Geriausiai tinka tolygusis skirstinys.
```

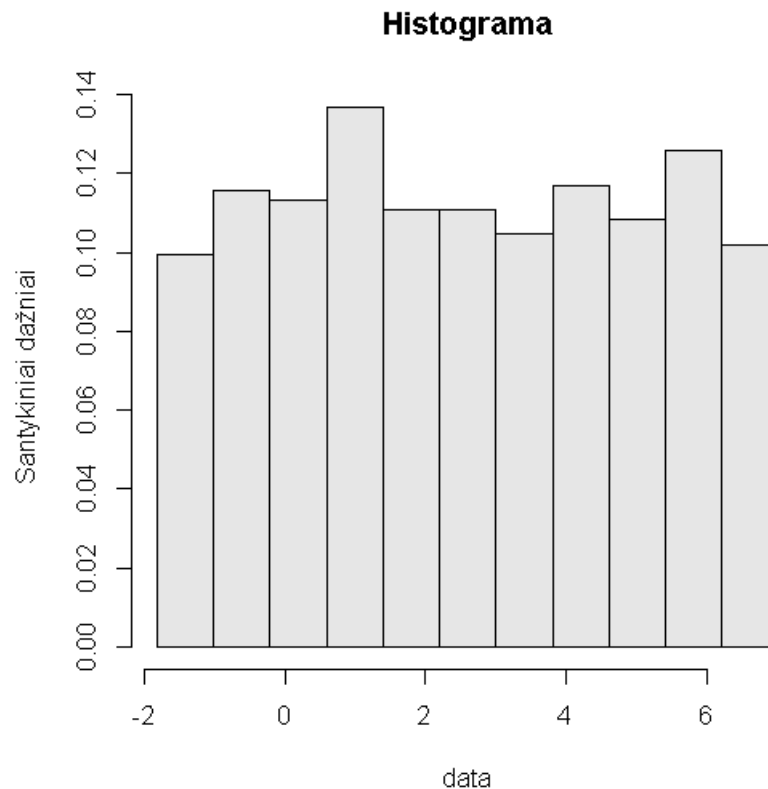
4 pav. 1.2. uždavinio R kodas

## REZULTATAI IR IŠVADOS

**Stačiakampė diagrama**



5 pav. 1.2. uždavinio pirmasis rezultatas



6 pav. 1.2. uždavinio antrasis rezultatas

Remdamiesi histograma, nustatome, kad mums tinka tolygusis skirstinys.

### 1.3. R PROGRAMOS KODAS

```
#geriausiai tinka tolygusis skirstinys.
#1.3. Raskite skirstinio parametru taškinius iverčius.
#Taškiniai iverčiai - didžiausio tiketumo metodas
install.packages("MASS")
library(MASS)
task_iverciai <- fitdistr(data, "normal")
task_iverciai
```

```
#geriausiai tinka tolygusis skirstinys.
#1.3. Raskite skirstinio parametru taškinius iverčius.
#Taškiniai iverčiai - didžiausio tiketumo metodas
#Tolygaus skirstinio taškiniai iverčiai

a = floor(min(data))
b = ceiling(max(data))
a
b
k <- ceiling(log10(length(data))*3.32+1)
k
bin = seq(a, b, length.out = k+1)
bin
k = k - 1
bin1 = c(a, bin[2:k], b)
bin1
p = c(punif(bin1[1:k+1],a,b) - punif(bin1[0:k],a,b))
p
sum(p)
p
```

7 pav. 1.3. uždavinio R kodas

## REZULTATAI IR IŠVADOS

```
> library(MASS)
> task_iverciai <- fitdistr(data, "normal")
> task_iverciai
      mean      sd
2.59408069 2.51991804
(0.07968681) (0.05634708)
```

```
> a = floor(min(data))
> b = ceiling(max(data))
> a
[1] -2
> b
[1] 8
> k <- ceiling(log10(length(data))*3.32+1)
> k
[1] 11
> bin = seq(a, b, length.out = k+1)
> bin
[1] -2.0000000 -1.0909091 -0.1818182  0.7272727  1.6363636  2.5454545  3.4545455  4.3636364  5.2727273  6.1818182  7.0909091  8.0000000
> k = k - 1
> bin1 = c(a, bin[2:k], b)
> bin1
[1] -2.0000000 -1.0909091 -0.1818182  0.7272727  1.6363636  2.5454545  3.4545455  4.3636364  5.2727273  6.1818182  8.0000000
> p = c(punif(bin1[1:k+1],a,b) - punif(bin1[0:k],a,b))
> p
[1] 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.18181818
> sum(p)
[1] 1
> p
[1] 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.09090909 0.18181818
```

8 pav. 1.3. uždavinio rezultatai

## 1.4 R PROGRAMOS KODAS

```
#1.4. Suformuluokite ir patikrinkite atitinkama suderinamumo hipotezę
 #(reikšmingumo lygmenį  $\alpha$  pasirinkite patys).
#Urašykite nagrinėjamo populiacijos požymio skirstinio funkciją.
#Taikydami Kolmogorovo kriterijų patikrinkite hipotezę,
#kad duomenys pasiskirstė pagal tolygųį dėsnį
a = -2
b = 7.2
# Surašome duomenis į variacinę eilutę:
data = sort(data)
data
n = length(data)
# Apskaičiuojame pasiskirstymo funkcijos reikšmes:
FoX = punif(data, a, b)
FoX
# Tarpiniai skaičiavimai Kolmogorovo statistikai D_n apskaičiuoti:
FnX = seq(1:n)/n # i/n
Fn1X = (seq(1:n)-1)/n # (i-1)/n
# Apskaičiuojame skirtumus su teorine funkcija.
Dp = FnX - FoX
Dm = FoX - Fn1X
# Suformuojame tarpinių rezultatų lentelę:
Lentele = cbind(data, FnX, Fn1X, FoX, Dp, Dm)
Lentele
# Iš lentelės išrenkame D_n+ ir D_n- reikšmes DP ir DM:
DP = max(Dp)
DP
DM = max(Dm)
DM
# Apskaičiuojame Kolmogorovo statistikos D_n reikšmę Dn:
Dn = max(DP,DM)
Dn
# Apibrėžiame reikšmingumo lygmenį alpha.
alpha = 0.05
# Pagal apytiksle formulę apskaičiuojame Kolmogorovo skirstinio kvantilį
k = sqrt(-log(alpha/2)/(2*n))-1/(6*n)
k
#Kadangi 0.03022493 < 0.04278027, galime teigti, kad atsitiktinis dydis
# pasiskirstė pagal tolygųį dėsnį, neprieštarauja imties duomenims.
#SKIRSTINIO FUNKCIJA ?????
#f(x) = ??
dnorm(data, emp_vidurkis, standartas_pat)
```

9 pav. 1.4. uždavinio R kodas

## REZULTATAI IR IŠVADOS

```
> # Apskaičiuojame kolmogorovo statistikos D_n reikšmę Dn:
> Dn = max(DP,DM)
> Dn
[1] 0.03022493
> # Apibrėžiame reikšmingumo lygmenį alpha.
> alpha = 0.05
> # Pagal apytikslę formulę apskaičiuojame kolmogorovo skirstinio kvantilį k:
> k = sqrt(-log(alpha/2)/(2*n))-1/(6*n)
> k
[1] 0.04278027
>
```

```
> dnorm(data, emp_vidurkis, standartas_pat)
[1] 0.12896415 0.05708753 0.05055012 0.08422118 0.13199689 0.05122218 0.12146194 0.04694708 0.12220882 0.15820077 0.15820015 0.12707172 0.07522796 0.13491224
[15] 0.15741040 0.12966120 0.15822845 0.15796870 0.13959312 0.08441706 0.04438773 0.12138648 0.10525208 0.10289899 0.13618743 0.07152778 0.15767822 0.14398821
[29] 0.15813894 0.11699331 0.11030733 0.09101196 0.04173076 0.13319290 0.03416361 0.09401712 0.07763285 0.06245375 0.09156593 0.05299224 0.05747836 0.09390671
[43] 0.04459184 0.14154506 0.15138810 0.12960847 0.15801439 0.14957580 0.15789144 0.10908873 0.13948747 0.15823461 0.07922203 0.08357154 0.13405116 0.05076059
[57] 0.14865589 0.04417911 0.15549404 0.11858338 0.03504166 0.06057993 0.09195656 0.15815408 0.15823037 0.14519337 0.14021499 0.15747564 0.15083316 0.13526385
[71] 0.03813113 0.15778938 0.10731911 0.13190127 0.14532301 0.13353693 0.09663410 0.07848243 0.09054511 0.09345787 0.13366835 0.08786290 0.11557480 0.03822167
[85] 0.04234242 0.11256346 0.06243859 0.07203057 0.14634875 0.06439557 0.15308891 0.07339713 0.09535157 0.11485738 0.15397869 0.15469489 0.15768114 0.15474613
[99] 0.03629770 0.11504776 0.07222789 0.09192576 0.15278087 0.09901284 0.14807719 0.06576523 0.08960068 0.14431425 0.10425626 0.09645805 0.04119527 0.03839100
[113] 0.15343674 0.07679679 0.10567948 0.14890914 0.06480507 0.04847505 0.13309436 0.12021895 0.04598350 0.15678567 0.11839947 0.13788671 0.11475324 0.13833109
[127] 0.12857688 0.09856595 0.11697526 0.09432346 0.07484126 0.03479356 0.06877873 0.12498584 0.14844824 0.11575233 0.14146188 0.14697002 0.06812824 0.11003246
[141] 0.15716224 0.04289098 0.13995804 0.15740160 0.07803632 0.06870022 0.04221060 0.15128967 0.07071005 0.12410216 0.12276887 0.10016267 0.15630402 0.03419607
[155] 0.13724260 0.06008309 0.15665934 0.06231540 0.04081018 0.11689740 0.10921621 0.03943261 0.13763857 0.04955175 0.09749659 0.08731050 0.09627041 0.14209509
[169] 0.12087839 0.13880907 0.10842049 0.07276470 0.14578729 0.09275132 0.12315216 0.0383734 0.15681062 0.11834600 0.14687815 0.05884532 0.13732543 0.06920437
[183] 0.15304346 0.06738777 0.08761057 0.03409066 0.15808114 0.15795659 0.14599773 0.12660730 0.03931148 0.07134955 0.14918376 0.13330379 0.05695584 0.13537591
[197] 0.14721128 0.04927260 0.15240481 0.15454167 0.13519348 0.15823514 0.14673577 0.13922278 0.15360297 0.05627094 0.15775822 0.06208540 0.11850639 0.11630942
[211] 0.15955154 0.15800730 0.15504280 0.11876070 0.12764240 0.07002721 0.15614498 0.04745896 0.15700442 0.12993036 0.14055346 0.12224935 0.14358919 0.12616584
[225] 0.11689229 0.11281947 0.06983417 0.14362061 0.14621743 0.09936990 0.12935729 0.14222502 0.10376222 0.10298533 0.08040369 0.04746520 0.05891011 0.11053333
[239] 0.13922828 0.15790232 0.06589238 0.15814710 0.15822557 0.03704877 0.14886893 0.08817112 0.09365527 0.15682663 0.12921887 0.06842033 0.12506140 0.15620773
[253] 0.04697752 0.14751438 0.15802162 0.07643672 0.10469004 0.09548350 0.12568301 0.04920849 0.15054404 0.12944531 0.12944467 0.10827665 0.06325620 0.15165033
[267] 0.13228007 0.15770803 0.04589272 0.15681036 0.06802108 0.10301718 0.04974654 0.07589463 0.12888204 0.07399867 0.14195072 0.12479662 0.14155991 0.14603721
[281] 0.12587103 0.12492895 0.15823538 0.10295397 0.07611653 0.15698304 0.03550804 0.07214224 0.15589940 0.13957093 0.04955175 0.15376758 0.15092109 0.15312959 0.15412783
[295] 0.08125982 0.13928655 0.06568983 0.15797394 0.15366080 0.08105799 0.06465303 0.10113310 0.08549673 0.15792499 0.15804616 0.15823113 0.09501169 0.10992601
[309] 0.0539872 0.05534872 0.08780200 0.05869758 0.14101438 0.15388633 0.13352799 0.15766379 0.14082705 0.14713777 0.14424828 0.12209348 0.04164082 0.07298601
[323] 0.03940821 0.13355206 0.07776107 0.14392582 0.11156311 0.06922387 0.14057073 0.09850543 0.13940449 0.10503516 0.15188159 0.04406438 0.09996885 0.04179563
[337] 0.12716843 0.14766372 0.05427376 0.08083401 0.11551505 0.08001414 0.08736852 0.08000105 0.06947489 0.05364939 0.06004024 0.05477356 0.10374154 0.15263822
[351] 0.08972886 0.09708486 0.11988617 0.03954046 0.13144161 0.13349868 0.11290939 0.05678392 0.05874656 0.15804327 0.14406116 0.14712682 0.07413073 0.14681940
[365] 0.06103199 0.15000972 0.05778366 0.11618656 0.15343355 0.10729226 0.03720077 0.09570161 0.08991711 0.03746830 0.15407077 0.05683802 0.14480195 0.12694864
[379] 0.14314605 0.05177934 0.13471595 0.09698513 0.13967830 0.07750149 0.15692144 0.03914292 0.05145234 0.15559497 0.14942163 0.12404984 0.13112287 0.15821969
[393] 0.11327181 0.15713031 0.15527595 0.03809802 0.14233828 0.12674980 0.10918915 0.03759479 0.10105596 0.09446785 0.11608260 0.06686120 0.07574655 0.07375448
[407] 0.04850564 0.07597196 0.13112879 0.07957991 0.14454183 0.15810152 0.13094559 0.14955595 0.15774441 0.13331746 0.15801354 0.10652769 0.07833584 0.15655671
[421] 0.11719542 0.14623783 0.08671681 0.06193790 0.08019281 0.15766417 0.14305449 0.13514570 0.10808922 0.11992901 0.15450235 0.14608558 0.15190864 0.05601936
[435] 0.09918139 0.08363591 0.11210235 0.14683595 0.05167035 0.13118563 0.13499095 0.08627918 0.05874018 0.04851127 0.07482437 0.09083935 0.15303175 0.13692013
[449] 0.08736451 0.13963317 0.14571374 0.15803329 0.15518879 0.05455634 0.06639772 0.03549435 0.12845924 0.06640167 0.04282114 0.15206402 0.15801394 0.08098852
[463] 0.07468257 0.15743165 0.15539811 0.06629833 0.13213272 0.08770552 0.13029843 0.15199943 0.04451382 0.11420854 0.15225850 0.12070585 0.04035561 0.03741063
[477] 0.10688398 0.06776100 0.11343845 0.13871957 0.04634628 0.13648090 0.05847731 0.05413580 0.05046583 0.05485623 0.11171261 0.11156333 0.13697878 0.04432727
[491] 0.14889391 0.05180683 0.09697588 0.08428400 0.15099666 0.09524972 0.15179678 0.04669727 0.11765298 0.09180863 0.04714697 0.15681218 0.12863395 0.07372222
[505] 0.11798155 0.11772243 0.12248936 0.10776209 0.08624953 0.14765881 0.15179633 0.04676278 0.07520897 0.08341744 0.13325209 0.11752724 0.15716168 0.05108445
[519] 0.10861384 0.14884735 0.07778401 0.15437459 0.14170486 0.14168471 0.08920730 0.12900176 0.10946158 0.11863901 0.06247955 0.09328617 0.09057269 0.08215810
[533] 0.13360203 0.09369472 0.03743020 0.05215417 0.11203203 0.15421698 0.08218348 0.15814602 0.04279277 0.04052622 0.10607596 0.03581941 0.14358250 0.11076249
[547] 0.10497293 0.14877442 0.13716872 0.14915996 0.12061251 0.12443684 0.04243878 0.14451353 0.11751598 0.14437330 0.07206442 0.10102685 0.07287686 0.15009856
[561] 0.14793543 0.14878666 0.09337920 0.15793142 0.04401845 0.11354635 0.05881419 0.11598653 0.15650052 0.05333111 0.15064069 0.14100536 0.11590528 0.13349499
[575] 0.03516075 0.14245316 0.09220275 0.09825656 0.07333209 0.15555297 0.12426253 0.14848012 0.09399356 0.12747256 0.11056577 0.05212194 0.12144393 0.15618274
[589] 0.05374225 0.13604005 0.11561931 0.07275223 0.07207211 0.04724301 0.14467290 0.15558319 0.07812418 0.13717612 0.06754575 0.15819681 0.15598760 0.04781752
[603] 0.15531661 0.04905847 0.13790504 0.06441174 0.07193679 0.15821789 0.13236553 0.15743634 0.04870236 0.14083312 0.15758910 0.05350237 0.05741969 0.03725872
[617] 0.13394663 0.06917767 0.05080163 0.14621845 0.11415014 0.05123801 0.06351533 0.07909445 0.04604275 0.13224576 0.07580789 0.07199700 0.05004776 0.08743269
[631] 0.15807603 0.15158955 0.13310030 0.04419428 0.14612141 0.14408352 0.09589395 0.06386225 0.04536543 0.04445131 0.11267852 0.05833509 0.05147250 0.08343769
[645] 0.15698731 0.07306375 0.08052346 0.12304307 0.04045211 0.13655321 0.12766116 0.05037956 0.04089739 0.07568473 0.04282227 0.12456063 0.11651202 0.07808542
[659] 0.09751377 0.12981161 0.14972686 0.10415813 0.04564424 0.09924283 0.12839135 0.06859631 0.07613637 0.06250988 0.03536627 0.11612899 0.09344812 0.05309932
[673] 0.09068956 0.08534579 0.06680201 0.07217158 0.14011635 0.06787506 0.15650941 0.10725080 0.13221832 0.07135350 0.08352926 0.07616487 0.09379064 0.15615842
[687] 0.15385684 0.07422362 0.09628390 0.06921977 0.12971407 0.10654735 0.06647289 0.09510390 0.09546934 0.04373428 0.15705703 0.05500698 0.05200223 0.04687165
[701] 0.04939836 0.03641437 0.11580072 0.06953431 0.13381092 0.13757932 0.14557412 0.06909407 0.08664209 0.11008328 0.14256275 0.08337525 0.13980959 0.09679635
[715] 0.04258687 0.07050164 0.15130985 0.11980627 0.11077832 0.12602437 0.03794349 0.05143157 0.09931108 0.15320577 0.15612965 0.06067535 0.10156703 0.13544667
[729] 0.13827163 0.05772861 0.11620084 0.12049307 0.09377760 0.09419409 0.05860667 0.14930546 0.06192388 0.15510697 0.12826371 0.05808340 0.14652204 0.11208184
[743] 0.15455132 0.13606886 0.07269282 0.05614026 0.07164308 0.15736816 0.15822784 0.15650312 0.04494213 0.15543243 0.11423608 0.10426440 0.14903674 0.03318865
```

10 pav. 1.4. uždavinio rezultatai

Iš 9-10 paveikslėlių rezultatų, tikrindami ar duomenys pasiskirste pagal tolydųjį dėsnį (Kolmogorovo dėsnis), padarome hipotezę (reikšmingumo lygmenį pasirenkame patys).



## 2. UŽDUOTIS.

### 2. GRAFINĖ ANALIZĖ IR PARAMETRINĖS HIPOTEZĖS

Iš duomenų imties „normalusis\_parametru\_iverciamis.csv“ nuskaitykite savo varianto duomenis tokiu būdu (nagrinėsime tik normaliuosius atsitiktinius dyžius):

```
N=variantas # Reikia įrašyti savo varianto numerį N
data=read.csv('normalusis_parametru_iverciamis.csv',header=TRUE) # Failas turi būti darbiname kataloge
pasiklovimo_lygmuo=data[1,N] # variantui skirtas pasiklovimo lygmuo  $\gamma=1-\alpha$ .
data=data[2:1001,N] # Duomenys, kuriuos reikia analizuoti - imtis
```

**2.1.** Nubraižykite histogramą, stačiakampę diagramą ir kvantilinį grafiką (Q-Q plot). Parašykite, kokios šių grafikų savybės rodo, kad nagrinėjamo atsitiktinio dydžio skirstinys yra normalusis (po vieną savybę iš kiekvieno grafiko).

**2.2.** Raskite taškinis vidurkio ir dispersijos įverčius.

**2.3.** Sudarykite parametų pasikliautinuosius intervalus su pasiklovimo lygmeniu  $\gamma=1-\alpha$ . Parašykite po vieną išvadą apie populiacijos parametro intervalinį įvertį.

**2.4.** Suformuluokite ir patikrinkite hipotezę apie vidurkio lygybę skaičiui  $m_0$  su alternatyviaja hipoteze:

- 1)  $H_a : m \neq m_0$  ir reikšmingumo lygmeniu  $\alpha$ , kai varianto numeris yra lyginis skaičius;
- 2)  $H_a : m > m_0$  ir reikšmingumo lygmeniu  $\alpha$ , kai varianto numeris yra pirminis skaičius;
- 3)  $H_a : m < m_0$  ir reikšmingumo lygmeniu  $\alpha$ , kai varianto numeris yra nelyginis ir nepirminis.

11 pav. Antro uždavinio sąlyga

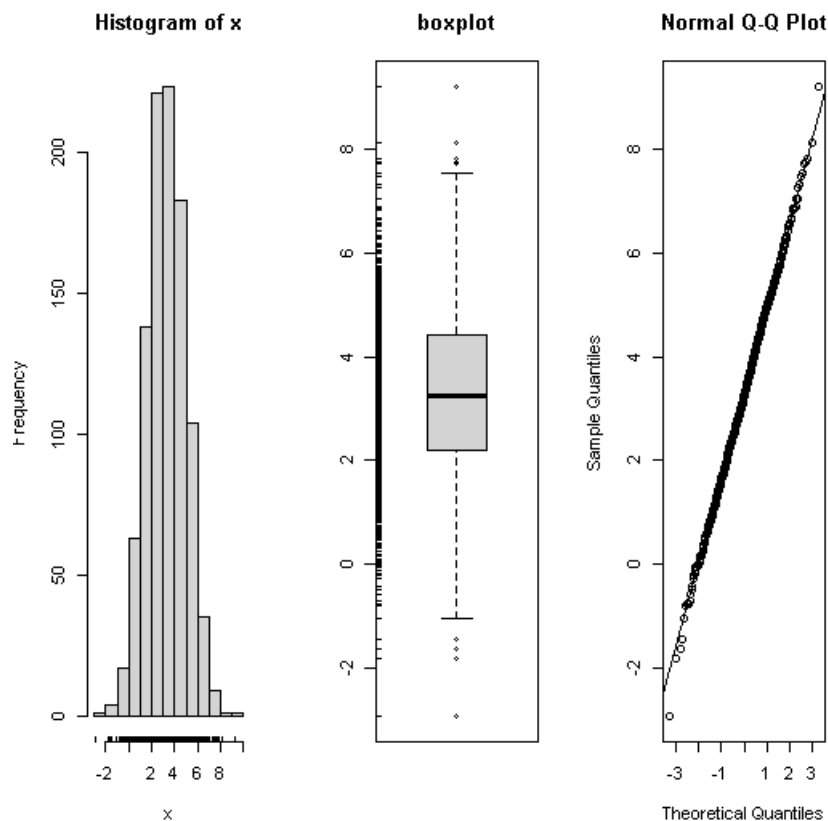
### 2.1. R PROGRAMOS KODAS

```
#2 DALIS
N=40 #varianto numeris
data=read.csv('normalusis_parametru_iverciamis.csv', header=TRUE)
pasiklovimo_lygmuo = data[1,N]
data = data[2:1001,N]
data
#2.1. Nubraižykite histogramą, stačiakampę diagramą ir kvantilinį grafiką (Q-Q plot).
#Parašykite, kokios šių grafikų savybės rodo,
#kad nagrinėjamo atsitiktinio dydžio skirstinys yra normalusis (po vieną savybę iš kiekvieno grafiko).
#vidurkio, modos ir medianos reikšmės sutampa
boxplot(data, horizontal = TRUE, main = "Stačiakampė diagrama")

k <- ceiling(log10(length(data))*3.32+1)
k
#skirstinio tikimybų pasiskirstymo kreivė yra dvipusiai simetriška, o simetrijos ašis yra ties vidurkiu
#Histograma
hist(data, breaks = seq(min(data), max(data),length = k+1),
      probability = T,
      right = F,
      col = grey(0.9),
      ylab = "Santykiniai dažniai",
      main = "Histograma")
)
##QQ diagramos išsidėstymas panašus į tiesę
library(UsingR)
simple.eda(data)
```

12 pav. 2.1. uždavinio kodas

## REZULTATAI IR IŠVADOS



13 pav. 2.1. uždavinio rezultatai

Išvados ir komentarai (12 paveikslėlis).

## 2.2. R PROGRAMOS KODAS

```
#2.2. Raskite taškinius vidurkio ir dispersijos įverčius.  
m=mean(data)  
m  
n=length(data)  
xvid = mean(data)  
xvid  
disp = ((n-1)/n)*var(data)  
install.packages("MASS")  
library(MASS)  
iverciai <- fitdistr(data,"normal")  
#vidurkio iverciai  
iverciai  
#dispersija  
disp
```

14 pav. 2.2. uždavinio R kodas

## REZULTATAI IR IŠVADOS

```
> library(MASS)  
> iverciai <- fitdistr(data,"normal")  
> #vidurkio iverciai  
> iverciai  
      mean      sd  
3.27111331 1.64547315  
(0.05203443) (0.03679390)  
> #dispersija  
> disp  
[1] 2.707582
```

15 pav. 2.2. uždavinio rezultatai

Vidurkio įverčiai (mean ir sd), dispersija (15 paveikslėlis).

## 2.3. R PROGRAMOS KODAS

```
#2.3. Sudarykite parametru pasisklidimusius intervalus su pasisklidimo lygmeniu  $\gamma=1-\alpha$ . Parašykite po vieną  
#išvadą apie populiacijos parametro intervalinį įvertį.  
#DISPERSIJA YRA ŽINOMA  
n = length(data)  
n  
xvid = mean(data)  
xvid  
pasisklidimo_lygmuo  
#kvantilis  
zp = qnorm(pasisklidimo_lygmuo)  
zp  
#Intervalinis įvertis (teorinės formules naudojimas)  
PI <- c(xvid-disp*zp/sqrt(n), xvid+disp*zp/sqrt(n))  
PI  
#Populiacijos parametro intervalas atitinka teorines reikšmes
```

16 pav. 2.3. uždavinio R kodas

## REZULTATAI IR IŠVADOS

```
> #2.3. Sudarykite parametru pasisklidimusius intervalus su pasisklidimo lygmeniu  $\gamma=1-\alpha$ . Parašykite po vieną  
> #išvadą apie populiacijos parametro intervalinį įvertį.  
> #DISPERSIJA YRA ŽINOMA  
> n = length(data)  
> n  
[1] 1000  
> xvid = mean(data)  
> xvid  
[1] 2.594081  
> pasisklidimo_lygmuo  
[1] 0.9  
> #kvantilis  
> zp = qnorm(pasisklidimo_lygmuo)  
> zp  
[1] 1.281552  
> #Intervalinis įvertis (teorinės formules naudojimas)  
> PI <- c(xvid-disp*zp/sqrt(n), xvid+disp*zp/sqrt(n))  
> PI  
[1] 2.484353 2.703809  
> #Populiacijos parametro intervalas atitinka teorines reikšmes
```

17 pav. 2.3. uždavinio rezultatai

Išvados ir rezultatai (16-17) paveikslėliuose, komentarai.

## 2.4. R PROGRAMOS KODAS

```
#2.4. Suformuluokite ir patikrinkite hipotezę apie vidurkio lygįje skaičių  $m_0$  su alternatyviąja hipoteze:  
#1)  $H_0: m = m_0$  ir reikšmingumo lygmeniu  $\alpha$ , kai varianto numeris yra lyginis skaičius;  
#2)  $H_0: m > m_0$  ir reikšmingumo lygmeniu  $\alpha$ , kai varianto numeris yra pirminis skaičius;  
#3)  $H_0: m < m_0$  ir reikšmingumo lygmeniu  $\alpha$ , kai varianto numeris yra nelyginis ir nepirminis.  
#1 variantas  $m_0 = 3.2$ , varianto numeris = 40  
#1 variantas - 40 yra lyginis skaičius.  
N=40 #varianto numeris  
data=read.csv('normalusis_parametru_iverciams.csv', header=TRUE)  
pasisklidimo_lygmuo = data[1,N]  
data = data[2:1001,N]  
data  
m0 = 3.2 #skaičius paimtas iš intervalinio įverčio reikšmių  
install.packages("UsingR")  
library(UsingR)  
library(datasets)  
data(msleep)  
attach(msleep)  
#p reikšmė yra mažesnė už reikšmingumo lygmenį,  $2.2e-16 < 0.95$ , todėl  
#galime atmesti nulinę hipotezę, kad vidutinė reikšmė yra lygi 3.2. Galime teigti jog ši reikšmė statistiškai skiriasi nuo 3.2  
t.test(sleep_total, mu = m0, conf.level = 0.95, alternative = 'two.sided')  
detach(msleep)
```

18 pav. 2.4. uždavinio R kodas

## REZULTATAI IR IŠVADOS

```
> library(datasets)  
> data(msleep)  
> attach(msleep)  
> #p reikšmė yra mažesnė už reikšmingumo lygmenį,  $2.2e-16 < 0.95$ , todėl  
> #galime atmesti nulinę hipotezę, kad vidutinė reikšmė yra lygi 3.2. Galime teigti jog ši reikšmė statistiškai skiriasi nuo 3.2  
> t.test(sleep_total, mu = m0, conf.level = 0.95, alternative = 'two.sided')  
  
One Sample t-test  
  
data: sleep_total  
t = 14.808, df = 82, p-value < 2.2e-16  
alternative hypothesis: true mean is not equal to 3.2  
95 percent confidence interval:  
 9.461972 11.405497  
sample estimates:  
mean of x  
 10.43373
```

Pagal gautus rezultatus, nustatome ar hipotezė teisinga (18-19) paveikslėliai. Išvados ir komentarai nuotraukuose.