6 Paskaita. Parametrinių hipotezių tikrinimas dviem imtims.

6.1 Dviejų normaliųjų populiacijų vidurkių ir dispersijų palyginimas.

Dažnai tenka lyginti dviejų normalių skirstinių vidurkius ir dispersijas. Išvados daromos remiantis atitinkamų imčių statistikų skirtumais. Jei skirtumai dideli, daroma išvada apie tai, kad populiacijų parametrai skiriasi. Svarbu skirti 2 atvejus – priklausomų ir nepriklausomų imčių atvejus. Nepriklausomų imčių pavyzdžiai:

1. Lyginami merginų ir vaikinų pažymių vidurkiai, ūgiai, svoriai, IQ, atlyginimai ir t.t.;
2. Dviejų skirtingų šalių gyventojų pajamų palyginimas;
3. Įvairiais aspektais lyginami miesto ir kaimo gyventojai;
4. Lyginama produkcijos, pagamintos Europoje ir Azijoje vidutinė darbo iki gedimo trukmė.

Priklausomų (porinių) imčių pavyzdžiai:

1. Tiriami tų pačių žmonių svoriai prieš dietą ir po jos;
2. Atliekamas tų pačių žmonių kraujo tyrimas prieš ir po vaistų naudojimo;
3. Tiriama, ar pardavėjų mokymas padidina parduotų prekių skaičių;
4. Tiriama, ar žmogus vienodai našiai dirba dieninėje ir naktinėje pamainose.

6.2 Dviejų nepriklausomų populiacijų vidurkių palyginimas. Populiacijų dispersijos lygios.

Dviejų populiacijų vidurkių palyginimui naudojami Stjudento kriterijai arba t-testai. Tarkime, turime atsitiktines imtis $(X_1, \ldots, X_n)$ ir $(Y_1, \ldots, Y_m)$, gautas stebint du nepriklausomus atsitiktinius dydžius $X \sim \mathcal{N}(\mu_X, \sigma^2), Y \sim \mathcal{N}(\mu_Y, \sigma^2)$, kurių vidurkiai μ_X, μ_Y nežinomi, abiejų dydžių dispersija yra ta pati ir nežinoma. Norime patikrinti hipotezę:

$$\begin{cases} H_0 : \mu_X = \mu_Y, \\ H_1 : \mu_X \neq \mu_Y. \end{cases}$$

Kriterijaus statistika

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}$$

turi Stjudento skirstinį su $n+m-2$ laisvės laipsniais, jei $\mu_X = \mu_Y$. Čia S_X^2, S_Y^2 – imčių dispersijos. Tegu reikšmingumo lygmuo lygus α. Kritinės sritys priklausomai nuo alternatyvos, pateiktos lentelėje.

1 lentelė: Alternatyvos ir kritinės sritys, kai $H_0 : \mu_X = \mu_Y$

Alternatyva H_1	H_0 atmetama	H_0 neatmetama
$\mu_X \neq \mu_Y$	$\lvert t\rvert > t_{\alpha/2}(n+m-2)$	$\lvert t\rvert \leq t_{\alpha/2}(n+m-2)$
$\mu_X > \mu_Y$	$t > t_\alpha(n+m-2)$	$t \leq t_\alpha(n+m-2)$
$\mu_X < \mu_Y$	$t < -t_\alpha(n+m-2)$	$t \geq -t_\alpha(n+m-2)$

Pavyzdys. Lygių teisių komisija tikrina, ar didelėje draudimo įmonėje nėra lyčių diskriminavimo mokant atlyginimus. Atsitiktinai parinkus 20 draudimo agentų ir 25 agentes, paaiškėjo, kad vidutinė agento alga, atskaičius mokesčius, yra 1300€, ($s_X = 200$), o agentės – 1155€, ($s_Y = 300$). Lygių teisių komisija nori žinoti, ar nepažeidžiamos agenčių teisės ($\alpha = 0,05$).

Sprendimas. Reikia patikrinti statistinę hipotezę

$$\begin{cases} H_0 : \mu_X = \mu_Y, \\ H_1 : \mu_X \neq \mu_Y. \end{cases}$$

Skaičiuojame statistikos realizaciją turimiems duomenims:

$$t = \frac{1300 - 1155}{\sqrt{19 \cdot 200^2 + 24 \cdot 300^2}} \sqrt{\frac{20 \cdot 25(20+25-2)}{20+25}} \approx 1,855.$$

Kritinė reikšmė $t_{\alpha/2}(n+m-2) = t_{0,025}(43) = 2,01$. Kritinė sritis $|t| > 2,01$. Kadangi $t = 1,855 \notin W$, tai H_0 neatmetama. Skirtumas tarp agentų ir agenčių vidutinių atlyginimų statistiškai nereikšmingas.

6.3 Dviejų nepriklausomų populiacijų vidurkių palyginimas. Populiacijų dispersijos nelygios.

Tarkime, turime atsitiktines imtis $(X_1, \ldots, X_n)$ ir $(Y_1, \ldots, Y_m)$, gautas stebint du nepriklausomus atsitiktinius dydžius $X \sim \mathcal{N}(\mu_X, \sigma_X^2), Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, kurių vidurkiai μ_X, μ_Y nežinomi, abiejų dydžių dispersijos nelygios. Norime patikrinti hipotezę:

$$\begin{cases} H_0 : \mu_X = \mu_Y, \\ H_1 : \mu_X \neq \mu_Y. \end{cases}$$

Kriterijaus statistika $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{S_X^2/n + S_Y^2/m}}$ turi Stjudento skirstinį su $k = \min(n-1, m-1)$ laisvės laipsniais, kai $H_0 : \mu_X = \mu_Y$ yra teisinga. Kritinės sritys nurodytoa 2-je lentelėje.

2 lentelė: Alternatyvos ir kritinės sritys, kai $H_0 : \mu_X = \mu_Y$

Alternatyva H_1	H_0 atmetama	H_0 neatmetama
$\mu_X \neq \mu_Y$	$\lvert t\rvert > t_{\alpha/2}(k)$	$\lvert t\rvert \leq t_{\alpha/2}(k)$
$\mu > \mu_Y$	$t > t_\alpha(k)$	$t \leq t_\alpha(k)$
$\mu < \mu_Y$	$t < -t_\alpha(k)$	$t \geq -t_\alpha(k2)$

Pavyzdys. Tirdami dviejų statybinių brigadų darbą, ekspertai matavo laiką, sugaištą nuo atvykimo į darbą iki darbo pradžios. Pirmosios brigados darbininkai buvo tikrinti 30 kartų. Jie vidutiniškai sugaišo 10,82 min.($s_X = 7,2$ min.), antrosios brigados darbininkai buvo tikrinti 32 kartus ir vidutiniškai sugaišo 9,02 min, ($s_Y = 1,2$), Ar duomenys patvirtina hipotezę, kad 2-oji brigada dirba geriau? $\alpha = 0,1$.

Sprendimas. Patikrinsime hipotezę apie dviejų vidurkių lygybę su vienpuse alternatyva:

$$\begin{cases} H_0 : \mu_X = \mu_Y, \\ H_1 : \mu_X > \mu_Y. \end{cases}$$

Hipotezei tikrinti skaičiuojame kriterijaus statistikos reikšmę:
$T = \dfrac{10,82 - 9,02}{\sqrt{7,2^2/30 + 1,2^2/32}} \approx 1,352$. Kritinė reikšmė $t_\alpha(k) = t_{0,1}(29) = 1,311$. Kritinė sritis $t > 1,311$. Kadangi $t = 1,352 > 1,311$, turime atmesti nulinę hipotezę. Taigi dviejų statybinių brigadų atvykimo į darbą laiko vidurkiai reikšmingai skiriasi – antroji brigada į darba atvyksta greičiau.

6.4 Hipotezė apie dviejų dispersijų lygybę. Nepriklausomos imtys.

Matome, kad prieš tikrinant hipotezę apie nepriklausomų populiacijų vidurkių lygybę, reikia patikrinti, ar lygios populiacijų dispersijos. Tarkime, turime atsitiktines imtis $(X_1, \dots, X_n)$ ir $(Y_1, \dots, Y_m)$, gautas stebint du nepriklausomus atsitiktinius dydžius $X \sim \mathcal{N}(\mu_X, \sigma_X^2), Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, kurių dispersijos nežinomos. Tikriname hipotezę:

$$\begin{cases} H_0 : \sigma_X^2 = \sigma_Y^2, \\ H_1 : \sigma_X^2 > \sigma_Y^2. \end{cases}$$

Kriterijaus statistika

$$F = \frac{S_X^2}{S_Y^2}, \text{ kai } S_X^2 > S_Y^2.$$

Kriterijaus statistika F turi Fišerio skirstinį su $n-1$ ir $m-1$ laisvės laipsniais, kai nulinė hipotezė yra teisinga. Praktikoje naudojama statistika
$F = \dfrac{\text{didesnioji dispersija}}{\text{mažesnioji dispersija}}$.

Tuomet F yra nemažesnis už 1. Jei n_1 yra imties su didesniąja dispersija elementų skaičius, n_2 yra imties su mažesniąja dispersija elementų skaičius, F turi Fišerio skirstinį su $n_1 - 1$ ir $n_2 - 1$ laisvės laipsniais, kai nulinė hipotezė yra teisinga. Kritinės sritys nurodytoa 3-je lentelėje.

3 lentelė: Alternatyvos ir kritinės sritys, kai $H_0 : \sigma_X^2 = \sigma_Y^2$

Alternatyva H_1	H_0 atmetama
$\sigma_X^2 \neq \sigma_Y^2$	$F < F_{1-\alpha/2}(n_1 - 1, n_2 - 1)$ arba $F > F_{\alpha/2}(n_1 - 1, n_2 - 1)$
$\sigma_X^2 > \sigma_Y^2$	$F > F_\alpha(n_1 - 1, n_2 - 1)$

Pavyzdys. Medicinos tyrėjas norėtų sužinoti, ar rūkančiųjų ir nerūkančiųjų širdies susitraukimų dažnio (širdies ritmo) dispersijos skiriasi. Duomenys yra pateikti 4-je lentelėje. Naudojant $\alpha = 0.05$ patikrinkite ar yra pakankamai įrodymų, kad būtų galima pagrįsti teiginį apie dispersijų lygybę?

4 lentelė: Empiriniai duomenys

Rūko	Nerūko
$n_1 = 26$	$n_2 = 18$
$s_1^2 = 36$	$s_2^2 = 10$

Sprendimas. Tikriname hipotezę:

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2, \\ H_1 : \sigma_1^2 > \sigma_2^2. \end{cases}$$

Kriterijaus statistika

$$F = \frac{S_1^2}{S_2^2} = \frac{36}{10} = 3,6.$$

Kritinė reikšmė $F_\alpha(n_1 - 1, n_2 - 1) = F_{0,05}(25, 17) = 2,18$. Kadangi $F = 3,6 > 2,18$, nulinė hipotezė atmetama, rūkančiųjų ir nerūkančiųjų širdies susitraukimų dažnio dispersijos skiriasi.

6.5 Dviejų priklausomų populiacijų vidurkių palyginimas.

Tarkime, kad turime imties duomenų poras:$(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, pvz., tų pačių studentų I ir II kurso žiemos sesijų rezultatų vidurkiai. Stebimi atsitiktiniai dydžiai $X \sim \mathcal{N}(\mu_X, \sigma_X^2), Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$. Nei vidurkiai μ_X, μ_Y, nei dispersijos σ_X^2, σ_Y^2 nėra žinomi. Norime patikrinti hipotezę:

$$\begin{cases} H_0 : \mu_X = \mu_Y, \\ H_1 : \mu_X \neq \mu_Y. \end{cases}$$

Randame porų duomenų skirtumus: $d_1 = x_1 - y_1, \ldots, d_n = x_n - y_n$. Kriterijaus statistika

$$t = \frac{\bar{d}}{\sqrt{s_d^2/n}}$$

turi Stjudento skirstinį su $n-1$ laisvės laipsniais, jei H_0 yra teisinga. Čia $\bar{d}$ ir s_d^2 yra skirtumų imties vidurkis ir dispersija, t.y.

$$\bar{d} = \frac{d_1 + d_2 + \ldots + d_n}{n}, s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2.$$

Tegu reikšmingumo lygmuo lygus α. Kritinės sritys priklausomai nuo alternatyvos, pateiktos 5-je lentelėje.

5 lentelė: Alternatyvos ir kritinės sritys, kai $H_0 : \mu_X = \mu_Y$

Alternatyva H_1	H_0 atmetama	H_0 neatmetama
$\mu_X \neq \alpha$	$\|t\| > t_{\alpha/2}(n-1)$	$\|t\| \leq t_{\alpha/2}(n-1)$
$\mu > \mu_Y$	$t > t_\alpha(n-1)$	$t \leq t_\alpha(n-1)$
$\mu < \mu_Y$	$t < -t_\alpha(n-1)$	$t \geq -t_\alpha(n-1)$

Pavyzdys. Dvylikai vyno ekspertų du kartus buvo pateiktas tas pats vynas. Pirmąkart – butelyje su prancūziška etikete, antrąkart – su rusiška. Ar galima teigti, kad etiketė turėjo įtakos vertinimams? $\alpha = 0,1$. Ekspertų vertinimai – poriniai duomenys (prancūziška etiketė, rusiška etiketė):

$(12, 10); (5, 1), (14, 11); (11, 10); (19, 14); (7, 6); (12, 12); (9, 3); (10, 6); (16, 17); (5, 6); (16, 12).$

Sprendimas. Tikriname hipotezę apie vidurkių lygybę su dvipuse alternatyva. Sudarome skirtumų imtį $2; 4; 3; 1; 5; 1; 0; 6; 4; -1; -1; 4$. Randame

$$\bar{d} = 2,333, s_d^2 = 5,52, t_{0,05}(11) = 1,796, t = \frac{2,333}{\sqrt{5,52/12}} = 3,44.$$

Kadangi $3,44 > 1,796$, H_0 atmetama. Etiketė turėjo įtakos vertinimams.

6.6 Parametrinių hipotezių tikrinimas dviem imtims paketo R pagalba.

Vidurkių palyginimui dvejose grupėse naudojama komanda **t.test**. Sugeneruokime po 10 realizacijų 2-jų kintamųjų X ir Y, turinčių standartinį normalųjį skirstinį. Tačiau, tai tik vienas eksperimentas. Norėdami užtikrinti rezultatų stabilumą, eksperimentą reikėtų pakartoti daug kartų. Kadangi kriterijaus statistika t turi Stjudento skirstinį su $n + m - 2 = 18$ laisvės laipsniais, galime grafiškai pavaizduoti Stjudento skirstinio su 18 laisvės laipsnių tankio funkciją

```
> x = rnorm(10)
> y = rnorm(10)
> t.test(x,y)

        Welch Two Sample t-test

data:  x and y
t = 0.88768, df = 17.285, p-value = 0.3869
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5485017  1.3470230
sample estimates:
mean of x mean of y 
0.5906256 0.1913650 
```

1 pav.: Dviejų normaliųjų skirstinių vidurkių palyginimas R. Dvipusė alternatyva.

```
> ttest = t.test(x,y)
> names(ttest)
[1] "statistic"   "parameter"   "p.value"     "conf.int"    "estimate"    "null.value" 
[7] "alternative" "method"      "data.name"  
> ttest$statistic
        t 
0.8876789 
> ts = replicate(1000,t.test(rnorm(10),rnorm(10))$statistic)
> mean(ts)
[1] 0.03310834
> range(ts)
[1] -3.486428  3.839122
```

2 pav.: Hipotezės tikrinimas 1000 kartų.

komandos **dt(pts,df=18)** pagalba bei tame pačiame grafike pavaizduoti statistikos empirines reikšmes (3 pav.)

```
pts = seq(-3.9,3.9,length=100)
plot(pts,dt(pts,df=18),col='red',type='l')
lines(density(ts))
```

Skirstinių sutapimą galima taip pat patikrinti quantile-quantile grafiko pagalba. Komanda **rt(1000,df=18)** generuoja Stjudento skirstinio su 18 laisvės laipsnių reikšmes (4 pav.)

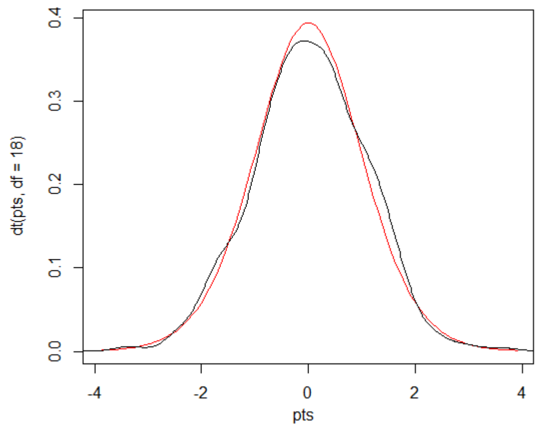
```
qqplot(ts,rt(1000,df=18))
abline(0,1)
```

Norėdami patikrinti hipotezę esant vienodoms dispersijoms, komanda t.test turi būti užrašyta taip:

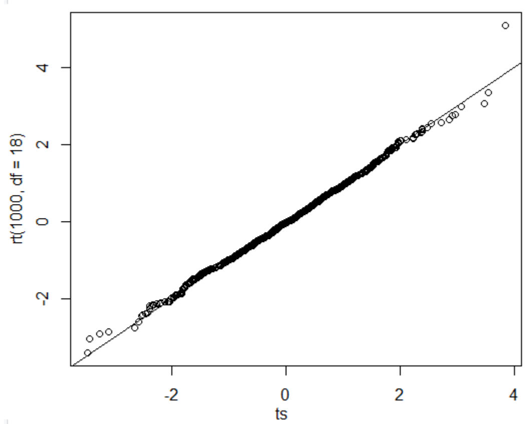
```
t.test(x,y,var.equal=TRUE)
```

Esant skirtingoms dispersijoms naudojama komanda

```
t.test(x,y,var.equal=FALSE)
```



3 pav.: Statistikos t empirinė ir teorinė tankio funkcijos.



4 pav.: Statistikos t empirinių ir teorinių reikšmių kvantiliai.

Dviejų populiacijų dispersijų lygybė tikrinama komandos **var.test** pagalba (5 pav.). Rezultatas – $p-$reikšmė yra pakankamai didelė (0,5487), be to dispersijų santykio pasikliautinasis intervalas apima reikšmę 1, t.y. dispersijos yra lygios.

Priklausomiems duomenims vidurkių palyginimui naudojama ta pati komanda **t.test** su opcija **paired = TRUE**. 6 pav. matome rezultatą, apskaičiuotą pavyzdiniems duomenims – peliukų svoriams prieš ir po eksperimento. Nulinė hipotezė atmetama, nes $p-$reikšmė ypatingai maža.

t.test(x, y, paired = TRUE, alternative = "two.sided")

```
> var.test(x,y)

        F test to compare two variances

data:  x and y
F = 0.66203, num df = 9, denom df = 9, p-value = 0.5487
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1644383 2.6653199
sample estimates:
ratio of variances
         0.6620278
```

5 pav.: Dviejų normaliųjų skirstinių dispersijų palyginimas R.

```
> # Weight of the mice before treatment
> before <-c(200.1, 190.9, 192.7, 213, 241.4, 196.9, 172.2, 185.5, 205.2, 193.7)
> # Weight of the mice after treatment
> after <-c(392.9, 393.2, 345.1, 393, 434, 427.9, 422, 383.9, 392.3, 352.2)
> t.test(before, after, paired = TRUE, alternative = "two.sided")

        Paired t-test

data:  before and after
t = -20.883, df = 9, p-value = 6.2e-09
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -215.5581 -173.4219
sample estimates:
mean of the differences
                -194.49
```

6 pav.: Dviejų priklausomų populiacijų vidurkių palyginimas R.