



**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**FUNDAMENTALIŲJŲ MOKSLŲ FAKULTETAS**  
**TAIKOMOSIOS MATEMATIKOS KATEDRA**

**Edita Besasparytė**

**MUZIKINIO RITMO STATISTINĖ**  
**ANALIZĖ**

Magistro darbas

**Vadovas**  
**dr. T. Ruzgas**  
**Konsultantas**  
**prof. Dr. D. Kučinskas**

**KAUNAS, 2011**



**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**FUNDAMENTALIŲJŲ MOKSLŲ FAKULTETAS**  
**TAIKOMOSIOS MATEMATIKOS KATEDRA**

**TVIRTINU**  
**Katedros vedėjas**  
**doc.dr. N.Listopadskis**  
**2011 06 02**

**MUZIKINIO RITMO STATISTINĖ**  
**ANALIZĖ**

Taikomosios matematikos magistro baigiamasis darbas

**Vadovas**  
**dr. T. Ruzgas**  
**2011 06 01**

**Recenzentas**  
**Prof. dr. Kęstutis Dučinskas**  
**2011 06 01**

**Atliko**  
**FMMM 9 gr. stud.**  
**E. Besasparytė**  
**2011 05 30**

**KAUNAS, 2011**

## KVALIFIKACINĖ KOMISIJA

**Pirmininkas:** Leonas Saulis, profesorius (VGTU)

**Sekretorius:** Eimutis Valakevičius, docentas (KTU)

**Nariai:** Algimantas Jonas Aksomaitis, profesorius (KTU)

Vytautas Janilionis, docentas (KTU)

Vidmantas Povilas Pekarskas, profesorius (KTU)

Rimantas Rudzkis, habil. dr., vyriausiasis analitikas (DnB NORD Bankas)

Zenonas Navickas, profesorius (KTU)

Arūnas Barauskas, dr., vice-prezidentas projektams (UAB „Baltic Amadeus“)

**Besasparytė E. Statistical analysis of musical rhythms : Master's work in applied mathematics / supervisor dr. T. Ruzgas; Department of Applied mathematics, Faculty of Fundamental Sciences, Kaunas University of Technology. – Kaunas, 2011. – 77 p.**

## **SUMMARY**

Statistical methods are one of the most popular objects used in creating compositions of modern music. It is practiced not just in interpretations of performance and physical quality of sound. Generating musical compositions and working on qualitative musical record storage are other fields of application. On the other hand statistical methods are not so often used for comparative analysis. The comparative analysis of musical data gives objective results that could be important when dealing with plagiarism problem. Another relevant issue is restoration of missing parts of compositions. A few examples of statistical analysis research in music are presented in overview of methods. In most cases analysis of musical works uses distribution results, which are interpreted in terms of music. M. K. Čiurlionis compositions were first analyzed by D. Kučinskas and R. Ambrazevičius in means of rhythms' distribution. Rhythm pulsations are compared between Čiurlionis compositions and folk music collections. These rhythms are analyzed in this paper with additional compositions included.

Musical rhythm is one easily written part of musical notation. Some features of rhythm are predictable, but it wouldn't be enough in comparison examination. Moreover the features of rhythm depend on musical meter, which could separate a composition into different parts. Issues in applying statistical methods for different data sections are studied. The basic task of investigation was to create the program that prepares raw data for analysis and makes its comparative research. The procedures of SAS system were used for analysis of variance, nonparametric tests, clustering and correspondence analysis. To measure similarity different coefficient were used and compared. The results show that collection of methods that were employed can be used in searching similarities between groups of musical pieces. This comparative analysis system could be applied for musical style or composer identification. It could group musical composition into blocks and show its dominant features.

# TURINYS

ĮVADAS.....	8
<b>1 TEORINĖ DALIS .....</b>	<b>9</b>
1.1 STATISTINIŲ TYRIMŲ MUZIKOJE APŽVALGA .....	9
1.2 MUZIKINIO RITMO TYRIMO GALIMYBĖS .....	10
1.3 METODŲ IR TYRIMO ĮGYVENDINIMO SAS APLINKOJE APŽVALGA .....	10
1.3.1 STATISTINIŲ METODŲ APŽVALGA.....	10
1.3.2 TYRIMO ĮGYVENDINIMAS SAS APLINKOJE .....	16
1.4 TYRIMO TIKSLAI.....	17
<b>2 TIRIAMOJI DALIS.....</b>	<b>18</b>
2.1 DUOMENŲ STRUKTŪRA IR ANALIZUOJAMI PJŪVIAI.....	18
2.2 MUZIKINIO RITMO STATISTINĖS ANALIZĖS MODELIAI .....	20
2.2.1 APRAŠOMOSIOS CHARAKTERISTIKOS .....	20
2.2.2 NEPRIKLAUSOMUMO TYRIMO MODELIS.....	23
2.2.3 PANAŠUMO KOEFICIENTAI IR KLASTERINĖ ANALIZĖ.....	26
2.2.4 ATITIKTIES TYRIMO MODELIS .....	29
<b>3 PROGRAMINĖ REALIZACIJA IR INSTRUKCIJA VARTOTOJUI .....</b>	<b>34</b>
3.1 DUOMENŲ ĮKĖLIMAS IR APDOROJIMAS.....	35
3.2 LYGINAMOJI RITMO ANALIZĖ .....	37
<b>4 DISKUSIJOS .....</b>	<b>41</b>
4.1 PASISKIRSTYMO Palyginimas ir nepriklausomumo tyrimas .....	41
4.2 PANAŠUMO KOEFICIENTAI IR KLASTERINĖ ANALIZĖ .....	42
4.3 ATITIKTIES TYRIMAS .....	42
4.4 REZULTATŲ Palyginimas.....	43
<b>IŠVADOS .....</b>	<b>44</b>
<b>REKOMENDACIJOS .....</b>	<b>45</b>
<b>PADĖKOS.....</b>	<b>46</b>
<b>LITERATŪRA .....</b>	<b>47</b>
<b>PRIEDAS NR.1 PRADINIS PROGRAMOS KODAS IR SCL FAILAI.....</b>	<b>49</b>
<b>PRIEDAS NR.2 XML ŽEMĖLAPIS.....</b>	<b>52</b>
<b>PRIEDAS NR.3 XML FAILO PAVYZDYS .....</b>	<b>55</b>
<b>PRIEDAS NR.4 DUOMENŲ APDOROJIMO PROGRAMINIAI KODAI.....</b>	<b>58</b>
<b>PRIEDAS NR.5 MODELIŲ PROGRAMINIAI KODAI .....</b>	<b>65</b>
<b>PRIEDAS NR.6 TYRIMO REZULTATŲ PAVYZDŽIAI.....</b>	<b>69</b>
<b>PRIEDAS NR.7 DARBE NAUDOJAMI MUZIKINIAI TERMINAI.....</b>	<b>77</b>

## LENTELIŲ SĄRAŠAS

Lentelė 2.1 Naudojamų muzikinių parametrų sąrašas .....	18
Lentelė 2.2 Taktų kirčių kvantiliai .....	22
Lentelė 2.3 Dispersinės analizės rezultatai .....	24
Lentelė 2.4 Tjukio grupavimo lentelė .....	24
Lentelė 2.5 Wilcoxon rangų sumos.....	25
Lentelė 2.6 Neparametrinių kriterijų rezultatų palyginimas .....	26
Lentelė 2.7 Formos koeficiento reikšmės .....	26
Lentelė 2.8 KL-divergencija simfoninės poemos „Jūra“ palyginimui.....	28
Lentelė 2.9 Simetrinė J-divergencija simfoninės poemos „Jūra“ palyginimui.....	29
Lentelė 2.10 Inercija ir chi-kvadrato statistikos atitikties tyrimo dimensijoms .....	29
Lentelė 2.11 Aprašomosios statistikos stulpelių taškams .....	30
Lentelė 2.12 Dalinė inercija stulpelių taškams .....	30
Lentelė 2.13 Kosinusų kvadratai stulpelių taškams .....	31
Lentelė 3.1 Apdorotų vertikalios pjūvio duomenų pavyzdys.....	35

## PAVEIKSLŲ SĄRAŠAS

Pav. 1.1 Programos struktūra .....	16
Pav. 2.1 Muzikinio ritmo taktų kodavimas .....	19
Pav. 2.2 Taktų kirčių pasiskirstymas pagal grupes .....	21
Pav. 2.3 Taktų kirčių pasiskirstymas kiekvienoje kūrinių grupėje .....	22
Pav. 2.4 Taktų kirčių histograma .....	23
Pav. 2.5 Kūrinių grupių dendrograma pagal panašumo koeficientą .....	27
Pav. 2.6 Takto kirčių grupavimas pagal panašumo koeficientą.....	28
Pav. 2.7 Atitikties žemėlapis.....	32
Pav. 2.8 Daugialypės atitikties analizės žemėlapis .....	33
Pav. 3.1 Programos katalogo struktūra .....	34
Pav. 3.2 Metodų įgyvendinimo struktūra programoje .....	34
Pav. 3.3 Duomenų įkėlimo ir apdorojimo programos langas.....	36
Pav. 3.4 Lyginamosios Ritmo Analizės langas.....	38
Pav. 3.5 Klasterinės analizės skiltis Analizės lange .....	39
Pav. 3.6 Atitikties tyrimo skiltis Analizės lange .....	39

## ĮVADAS

Viena populiariausių sričių taikomų moderniosios muzikos kūrimui yra statistiniai metodai. Jų taikymas apima nuo atlikimo ir garso fizikinių savybių interpretavimo iki muzikos generavimo ir kokybiško saugojimo. Statistinių metodų taikymas lyginamojoje muzikos analizėje nėra taip plačiai naudojamas kaip naujų kūrinų komponavime. Tačiau muzikinių duomenų palyginimas ir tam tikrų jų savybių objektyvus išryškinimas yra svarbus sprendžiant plagiavimo problemą. Taip pat vienas aktualių klausimų – trūkstamų rankraščių, kompozicijų dalių atkūrimas.

M. K. Čiurlionio kūrybos analizė ir ritmų pasiskirstymo tyrimai yra pradėti D. Kučinsko ir R. Ambrazevičiaus. Tyrimuose lyginamos M. K. Čiurlionio ritmų pulsacijos, jų išskirtinumas lyginant su liaudies muzikos rinkiniais pasiskirstymo atžvilgiu. Šie kūriniai, kartu su papildomais, analizuojami šiame darbe aprašomuose tyrimuose taikant statistinius metodus.

Muzikinis ritmas vienas paprasčiausiai užrašomų muzikos elementų. Tam tikros jo savybės yra nuspėjamos, tačiau norint palyginti ritmo savybes skirtingose kūrinų grupėse, neužtenka pasiskirstymo grafikų. Be to ritmo savybės skiriasi ir muzikinio metro atžvilgiu, todėl vieno kūrinio skirtingo metro ritmai gali būti analizuojami kaip skirtingi kūriniai. Darbe pateikta metodų tinkamų ritmo formulių tyrimui apžvalga. Išanalizuojamos problemos su kuriomis susiduriama taikant metodus skirtingiems duomenų pjūviams. Tyrimui išskelti uždaviniai:

- sukurti duomenų įkėlimo ir paruošimo sistemą, skirtą neapdorotiems muzikiniams failams;
- duomenų paruošimo procesą pritaikyti tiek horizontaliam tiek vertikaliam duomenų pjūviams, kur analizuojami atskiri taktai, arba taktai padalinti į takto kirčius;
- iširti metodų taikymo galimybes lyginamajai analizei, atlikti jų patikimumo tyrimą;
- sukurti programinę įrangą, leidžiančią įgyvendinti tam tikrą lyginamųjų metodų rinkinį.

Pastebėta, kad patogiau palyginamajai analizei naudoti vertikalų duomenų pjūvį, kuriame sukuriama viso kūrinio ar kūrinų grupės ritmo pulsacijų vektoriai. Tokių duomenų analizės interpretavimui ir tyrimui su kiekvienu metodu pateikiama vaizdinė arba grafinė priemonė. Naudojant dispersinės, klasterinės ir atitikties analizės metodus, apskaičiuojant atstumus, panašumo koeficientus, gauti rezultatai gali būti palyginami ir apibendrinami bendra išvada. Ši analizės sistema galėtų būti taikoma kūrinio stiliaus ar kompozitoriaus atpažinimui, arba kūrinų ar jų grupių palyginimui, dominuojančių savybių atskleidimui, tiriamų duomenų kategorijų gretinimui.



# 1 TEORINĖ DALIS

## 1.1 STATISTINIŲ TYRIMŲ MUZIKOJE APŽVALGA

Populiariausi muzikoje atliekami tyrimai yra susiję su jos generavimu arba komponavimu. Taikomi fizikos, statistikos metodai, struktūros tokios kaip fraktalai, Fibonacci sekos, simetrijos elementai (Miranda 2002). „Stochastinė muzika“ vieno garsiausių muzikos teoretikų graiko Iano Chenakio (*Iannis Xenakis*) sukurtas terminas (Schoner 2003). Šis terminas apibūdina muzikos komponavimą paremtą stochastiniais procesais. Šio kompozitoriaus, muzikos teoretiko, architekto tyrimai ir sukurti kompozicijos būdai padarė didelę įtaką elektroninės muzikos vystymuisi. Kai kurie šių metodų, taikomų kompozicijos kūrimui, gali būti taikomi tirti esamus kūrinius.

Kita plati sritis apimanti ne tik komponavimą bet ir kompozicijų analizę yra garso bangų tyrimai. Šiuo atveju statistiniais metodais tiriami garso dažniai, muzikinės akustikos, kompozicijos atlikimo savybės. Remiantis genetiniais algoritmais, kuriami modeliai muzikinių įrašų perrašymui, siekiant geresnės polifoninių kūrinių įrašų kokybės (Lu 2006).

Jei žiūrėsime į kompozicijos tekstą – jo parametrai apibūdina visas muzikines savybes, tačiau jų interpretavimas yra kiek sudėtingesnis. Dėl šios priežasties kompozicijos teksto tyrimai dažniausiai atliekami muzikologijos teoretikų. Muzikologijos tyrimų rezultatų interpretavimas gali būti subjektyvus, tačiau pasirinkus tam tikras koduotes, ar duomenų transformacijas kompozicijų tyrimui gali būti taikomi matematiniai metodai. Pavyzdžiui Markovo grandinės taikytos Yi-Wen Liu muzikinių duomenų paruošimui. Markovo 1-os eilės grandinės, sudarančios dvimates tikimybių matricas, lygintos tarpusavy Kullback-Leibler atstumo metrikos pagalba (Liu 2006). Šiuo pagrindu sukurtas modelis kompozitoriaus atpažinimui.

Įvairių technikų yra taikyta ritmo analizei: nuo antropologinių metodų iki lingvistinės pasiskirstymo analizės (Dahlig-Turek 2009). E.Dahlig-Turk įgyvendino 791 kūrinio, kurių ritmai buvo užrašomi skaitmeninio kodo forma, kompiuterizuotą tyrimą. Atliektais skaičiavimais nustatė ritmo parametrus – vidutinį ritminį tankį ir mažėjančią kokybę, sukonstravo ritmines matricas leidžiančias nagrinėti ritminių formulių chronologiją (Dahlig-Turek 2009).

Lietuvoje tyrimams muzikoje statistinius metodus taikė R.Kašponis, R.Ambrazevičius, D.Kučinskas. R.Kašponis lietuvių liaudies muzikos tyrimuose lygino muzikinio teksto dydžių pasiskirstymą, požymių priklausomumui tirti naudojo chi-kvadrato statistiką (Kašponis 1992). D.Kučinskas ir R.Ambrazevičius M.K.Čiurlionio fortepijoninės ir liaudies kūrybos tyrimuose lygino taktų pulsacijų pasiskirstymą. Jų tyrimuose gauti rezultatai galėtų būti lyginami su šiame darbe atliekamų tyrimų rezultatais.

## 1.2 MUZIKINIO RITMO TYRIMO GALIMYBĖS

Ritmas muzikoje palaiko tikslumą, įdeda į tam tikrus rėmus. Jis turi skirtingas savybes priklausomai nuo muzikinio metro, kūrinius ir jų dalis būtų galima grupuoti pagal metrą, kuris nusako kiek natų yra viename takte. Matematikoj metru atitikmuo būtų trupmena, proporcija. Galime sakyti, kad metras suskirsto kūrinį į lygias dalis – taktus. Žiūrint į taktą kaip į atskirą elementą, jo ilgis priklauso nuo metro t.y. jame telpa tam tikras kiekis ketvirtinių natų, dvigubas kiekis aštuntinių, ar trigubas kiekis šešioliktinių natų. Tad jeigu norėtume užrašyti takto ritmą simboliškai, pavyzdžiui  $\frac{3}{4}$  metre šešioliktinės natos tikslumu, turėtume 12 elementų taktą, kurį vadinsime ritmo chromosoma, sudarytą iš vienetų ir nulių, kur 1 reiškia natos pradžia, 0 – tylą arba besitęsiančią natą. Taigi yra galimybė tirti tuos pačius taktus skirtingo ilgio natos tikslumu, tačiau kuo ilgesnės natos tikslumu tiriama, tuo daugiau duomenų atmetama. Jeigu yra netipinio ritmo intarpų, tokių kaip trelis arba papuošimas (angl. *grace*) – intarpai ignoruojami, nes jie neužima takto ilgio. Taktai kuriuose yra triolės ar daugiau sugrupuotų natų (angl. *tuplet*) neanalizuojami. Šių muzikinių terminų paaiškinimai atitinkamai šiam tyrimui pateikti priede Nr.7. Taip pat gali būti atmetami taktai kuriuose nėra natų, nes jie nesukuria ritmo ir neturi esmės, kai analizuojamos visos melodinės linijos atskirai.

## 1.3 METODŲ IR TYRIMO ĮGYVENDINIMO SAS APLINKOJE APŽVALGA

### 1.3.1 STATISTINIŲ METODŲ APŽVALGA

**Horizontalus duomenų pjūvis.** Analizuojamas dydis – taktas, tam tikro ilgio ritmo chromosoma. Šių duomenų pjūvio tyrime galėtų būti taikomos požymių priklausomumo lentelės, chi-kvadrato kriterijus, koreliacijos koeficientas, atstumų koeficientai ir klasterinė analizė.

**Vertikalus duomenų pjūvis.** Apskaičiavus dažnius kiekvienam takto kirčiui ir atlikus normavimą, gaunami kūrinių ar kūrinių grupių vektoriai, kurių koordinatės atitinka natos atsiradimo tam tikrame kirtyje tikimybes. Pradžiai gali būti pateikiamos turimų vektorių empirinės charakteristikos, pagal turimas duomenų grupes pasitelkiant grafines priemones. Išsamesniems tyrimams taikytina dispersinė analizė, panašumo koeficientai ir atstumai, koreliacinė analizė, atitikties tyrimas.

Chi-kvadrato kriterijus taikomas nepriklausomumo hipotezei tikrinti tokiems duomenims, kurių tikėtini dažniai visuose požymių priklausomumo lentelės langeliuose yra nemažesni už 1 ir mažesni už 5 tikėtinų dažnių yra ne daugiau kaip 20% langelių. Jei šios sąlygos netenkinamos, dalis duomenų atmetama, nes duomenų apjungimas nebūtų tikslingas muzikiniu aspektu.

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - E_{ij})^2}{E_{ij}}, \quad (1.1)$$

čia  $E_{ij} = \frac{n_{i \cdot} \cdot n_{\cdot j}}{n}$  - tikėtinių reikšmių įverčiai, o  $n_{ij}$  - stebimi dažniai.

Paprastiau interpretuojamas koeficientas, gaunamas iš ch-kvadrato – Kramerio  $V$  koeficientas:

$$V = \frac{\sqrt{\chi^2 / n}}{\sqrt{\min(I-1, J-1)}}, \quad 0 \leq V \leq 1. \quad (1.2)$$

Nors takto kirčiai žymimi skaitmenimis, jie atitinka vardų skalės elementus. Kiekvienas takto kirtis yra nepriklausomas vienas nuo kito, todėl tikimybių rinkinį galime traktuoti kaip vektorių, arba tašką daugiamatėje Euklido erdvėje (Cha 2007). Tokiems vektoriams gali būti taikomi ir įprastiniai ir specifiniai diskretiesiems pasiskirstymo tankiams lyginti skirti panašumo koeficientai. Vienas paprasčiausių – Minkovskio  $L_p$  šeimos koeficientas – Euklido atstumas (1.3). Lyginame du vektorius arba empirinius pasiskirstymo tankius  $x = (x_1, \dots, x_n)$  ir  $y = (y_1, \dots, y_n)$ , kur visų  $j$  elementų svoriai vienodi.

$$d(x, y) = \sqrt{\sum_{j=1}^n (x_j - y_j)^2} \quad (1.3)$$

Naudojamas atstumas palyginimui formos atžvilgiu – „formos“ atstumas (angl. *shape*):

$$d_s(x, y) = \sqrt{\sum_{j=1}^n ((x_j - \bar{x}) - (y_j - \bar{y}))^2} \quad (1.4)$$

Panašumo koeficientas (angl. *similarity ratio* arba *simratio*) – artimiausias „vidinės sandaugos“ šeimos koeficientams (angl. *Inner Product family*):

$$s(x, y) = \frac{\sum_{j=1}^n x_j y_j}{\sum_{j=1}^n x_j y_j + \sum_{j=1}^n (x_j - y_j)^2} \quad (1.5)$$

Entropijos šeimos Kullback-Leibler koeficientas, taikomas diskrečių tikimybių tankių palyginimui:

$$d_{KL}(x, y) = \sum_{j=1}^n x_j \log_2 \frac{x_j}{y_j}, \quad (1.6)$$

čia  $\sum_{j=1}^n x_j = 1$  ir  $\sum_{j=1}^n y_j = 1$ . Reikšmės  $x_j, y_j > 0$ , nes kitu atveju  $d_{KL} = 0$ .

Šis atstumas yra neneigiamas, nesimetrinis  $x$  ir  $y$  atžvilgiu, ir gali artėti į begalybę (Shlens 2007). Simetrinė šio atstumo forma yra vadinama Jeffreys – Kullback – Leibler atstumu, arba J-divergencija (Kumar 2006):

$$d_J(x, y) = d_{KL}(x, y) + d_{KL}(y, x) = \sum_{j=1}^n (x_j - y_j) \log_2 \frac{x_j}{y_j}, \quad (1.7)$$

Taip pat pasiskirstymams palyginti naudojamas empirinis koreliacijos koeficientas:

$$r(x, y) = \frac{\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{j=1}^n (x_j - \bar{x})^2 \sum_{j=1}^n (y_j - \bar{y})^2}} \quad (1.8)$$

Čia  $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$  ir  $\bar{y} = \frac{1}{n} \sum_{j=1}^n y_j$ .

Panašumo (*simratio*) ir koreliacijos koeficientas absoliutiniu didumu įgyja reikšmes tarp 0 ir 1, kur didesnė koeficiento reikšmė reiškia stipresnį ryšį ar panašumą. Tyrime išbandyti ir kiti panašumo koeficientai skirti intervalų ir proporcijų tipo duomenims palyginti. Siekiant palyginti duomenis kaip daugiamatės erdvės taškus (daugiamačius vektorius), pasirinktas Euklido atstumas, vertinant duomenų vektorius kaip empirinius pasiskirstymo tankius (histogramas), palyginimui naudojamas formos atstumas (*shape*) ir pateiktos divergencijos. Papildomai rezultatų palyginimui naudojamas koreliacijos koeficientas, panašumo koeficientas. Klasterinėje analizėje naudojami atstumai ir koeficientai transformuoti į atstumus, pavyzdžiui panašumo koeficientas paverstas atstumu  $d_s = 1 - s(x, y)$  yra priešingas dydžiui 1.5 formulėje (angl. sutr. *disratio*).

Klasterinėje analizėje tiriant nedidelę stebėjimų imtį dažnai naudojami hierarchiniai metodai. Taikant jungimo būdą stebėjimai priskiriami ir apjungiami į klasterius remiantis pasirinktu metodu (pvz. *ward*, *average*, *centroid*), apibrėžiančiu atstumą tarp klasterių arba stebėjimų porų. Pagrindinis šio grupavimo rezultatas – dendrograma arba grupavimo medis, suteikianti vartotojui galimybę pačiam pasirinkti kirtimo liniją, ir nuspręsti kiek grupių, arba kiek elementų vienoje grupėje turėtų būti. Galimi ir kiti klasterių skaičiaus nustatymo būdai, remiantis pateikiamų statistikų reikšmėmis (Yeo 2003).

Kita grupavimo vizualizacija gaunama parametrinės dispersinės analizės pagalba – Tjukio grupavimo lentelė. Dispersinė analizė tiria dydžių nepriklausomumą nuo pasirinktų parametru

analizuojamame modelyje, t.y. tikrinama hipotezė apie vidurkių lygybę  $H_0 : m_1 = m_2 = \dots = m$  naudojant Fišerio statistiką (Analysis of Variance):

$$F = \frac{\overline{SS}_A}{\overline{SS}_e}, \quad (1.9)$$

čia  $\overline{SS}_A$  - nuokrypų kvadratų sumos vidurkis, apibūdinantis faktoriaus  $A$  poveikį stebimo atsitiktinio dydžio vidurkiui,  $\overline{SS}_e$  - nuokrypų kvadratų sumos vidurkis, apibūdinantis atsitiktinių klaidų poveikį stebimo atsitiktinio dydžio vidurkiui.

Tačiau parametrinė dispersinė analizė remiasi prielaida, kad dydžiai pasiskirstę pagal normalųjį skirstinį (Pappas 2004). Tokiu atveju rezultatų interpretavimas gali būti netikslingas, todėl patariama taikyti neparametrinę dispersinę analizę.

Tarkim turime atsitiktines imtis kiekvienai lyginamai populiacijai:  $X_1, \dots, X_{n_1}$  ir  $Y_1, \dots, Y_{n_2}$ . Šių neparametrinių kriterijų taikymui turimi  $N$  stebiniai sustatomi didėjančia tvarka ir kiekvienam  $j$ -tajam stebiniui priskiriamas rangas  $R_j$  (Narayanan). Jei  $k$  stebinių sutampa, kiekvienam jų priskiriama  $k$  iš eilės einantys rangai, vadinasi turime  $N$  rangų. Remiantis šiais rangais kiekvienam stebiniui priskiriami koeficientai (angl. *scores*)  $w_j$ . Taikomi kelių tipų koeficientai:

Wilcoxon-Mann-Whitney koef.	$w_j = R_j$	(1.10)
Medianos koef.	$w_j = \begin{cases} 1, & \text{kai } R_j > (N+1)/2 \\ 0, & \text{kitur} \end{cases}$	(1.11)
Van der Waerden koef.	$w_j = \Phi^{-1}(R_j / (N+1))$ , čia $\Phi \sim N(0,1)$	(1.12)
„Savage“ koef.	$w_j = \sum_{i=1}^{R_j} \frac{1}{N-i+1} - 1$	(1.13)

Jei analizuojama daug imčių, t.y.  $N$  stebinių kurie suklasifikuoti į  $k > 2$  lygių, su  $n_i$  stebinių  $i$ -tajame lygyje, kiekvienam stebiniui  $X_{ij}$  turėtų būti priskiriamas rangas  $R_{ij}$ .

Naudojant Wilcoxon'o koeficientus daugelio imčių atveju įgyvendinamas Kruskal-Wallis kriterijus. Apskaičiuojama Kruskal-Wallis statistika

$$S = \left( \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} \right) - 3(N+1), \text{ kur } R_i = \sum_{j=1}^{n_i} R_{ij}, \quad (1.14)$$

Tuomet randama  $p$ -reikšmė, kuri įvertinama chi-kvadrato statistikos pagalba  $P(S \leq \chi_{k-1}^2)$ . Tačiau jei lygmenų imčių dydis yra mažas (mažesnis nei 5)  $S$  pasiskirstymas gali skirtis nuo chi-kvadrato pasiskirstymo.

Kitas modelis turimų duomenų grupių palyginimui – atitikties analizė (angl. *correspondence analysis*), dar vadinama atitikties atvaizdavimu, atitikties faktorių analize. Šis tyrimas yra pagrindinių komponentų analizės krypties tyrimas skirtas kategoriniams duomenims (Friendly 2001). Atitikties analizėje naudojamas chi-kvadrato atstumas vietoj Euklido atstumo tarp taškų (Garson 2008). Atitikties žemėlapi sudaro dvi dimensijos, kurios atitinka du dominuojančius faktorius. Dažniausiai pasirenkamos dvi dimensijos, nes kitu atveju grafinį vaizdavimą būtų sudėtinga interpretuoti, be to dažniausiai pirmos dvi dimensijos paaiškina didžiąją dalį duomenų taškų pasiskirstymo. Atitikties analizė yra neparametrinis metodas nereikalaujantis, kad duomenys atitiktų kokį nors skirstinį. Vienintelis apribojimas – tikimybių lentelės elementai negali būti neigiami. Taikant šį metodą priimama prielaida apie stulpelių ir kategorijų eilučių homogeniškumą.

Tarkim tikimybių arba dažnių lentelė yra  $(I \times J)$  dydžio matrica, sudaryta iš neneigiamų elementų  $n_{ij}$ . Eilutės ir stulpeliai atitinka tam tikras kategorijas arba charakteristikas. Šios lentelės nuokrypis apskaičiuojamas chi-kvadrato statistikos pagalba (1.1). Parametras inercija (angl. *inertia*) yra gaunama chi-kvadrato statistiką padalinus iš visų dažnių lentelėje sumos (Micheloud 1997):

$$\phi^2 = \frac{\chi^2}{n} \quad (1.15)$$

Kuo didesnė inercija, tuo didesnis ryšys tarp eilučių ir stulpelių.

Chi-kvadratas gali būti užrašomas taip

$$\text{tr}(M^T R M C), \quad (1.16)$$

kur  $R = \text{diag}(1/n_{i\bullet})$ ,  $C = \text{diag}(1/n_{\bullet j})$ , o  $M = (n_{ij} - n_{i\bullet}n_{\bullet j}/n)$ .

Atitikties analizė susideda iš singuliarios reikšmės skaidinio (angl. *singular value decomposition* SVD) paieškos matricai  $R^{1/2} M C^{1/2}$  (Benko). Remiantis šiuo skaidiniu randamos eilučių ir stulpelių koordinatės. Tuo tarpu  $\chi^2$  - atstumai tarp eilučių arba stulpelių yra apskaičiuojami formulėmis (Greenacre):

$$\chi^2 \text{ tarp } i \text{ ir } i' \text{ eilučių: } \sum_{j=1}^J \left( \frac{n_{ij}}{n_{i\bullet}} - \frac{n_{i'j}}{n_{i'\bullet}} \right)^2 / n_{\bullet j} \quad (1.17)$$

$$\chi^2 \text{ tarp } j \text{ ir } j' \text{ stulpelių: } \sum_{i=1}^I \left( \frac{n_{ij}}{n_{\bullet j}} - \frac{n_{ij'}}{n_{\bullet j'}} \right)^2 / n_{i\bullet} \quad (1.18)$$

Inerciją dar galima apskaičiuoti turint singuliaras reikšmes, t.y. tikrines reikšmes:

$$inertia = \sum_{k=1}^K \alpha_k^2 = \sum_{k=1}^K \lambda_k, K = \min\{I-1, J-1\} \quad (1.19)$$

Atitikties žemėlapyje interpretavimas. Atitikties analizė pateikia visas tiriamas duomenų kategorijas Euklido erdvėje. Pirmosios dvi dimensijos pateikiamos ryšio tarp kategorijų ištyrimui. Jei imtume mažiausią tikimybių lentelės dydį sudarytą iš dviejų kategorinių eilučių ir vieno stulpelio, plokštumoje atvaizduoti taškai galėtų būti interpretuojami kaip du atskiri grafai kiekvienam kategoriniam kintamajam (SAS Institute 1999). Šiuo atveju būtų tiriamas ryšys tarp kategorinių kintamųjų esančių eilutėse su stulpelių kintamaisiais. Tačiau šiuo atveju kiekvienoje plokštumoje esantys taškai negali būti lyginami atstumo atžvilgiu. Atstumas tarp taškų turi prasmę tik kintamojo (eilučių arba stulpelių atžvilgiu) viduje tarp kategorijų, kadangi eilutės ir stulpeliai gali turėti skirtingą mastelį (Micheloud 1997). Kintamojo viduje palyginimas gali būti įgyvendinamas ir ašių arba faktorių pagalba: kuo mažesnis kampas su ašimi, tuo didesnis koreliuotumas (artimesnis 1), kuo kampas artimesnis  $90^\circ$  – tuo koreliacija artimesnė 0. Vienas pagrindinių atitikties analizės rezultatų – inercija. Bendroji inercija apsprendžia kokią dalį taškų atitikties žemėlapyje paaiškina visos dimensijos arba faktoriai, kiekvienos dimensijos inercija – atskirai kiekvienos ašies reikšmingumą. Taip pat kiekvienas taškas užima dalį bendrosios visų taškų žemėlapyje inercijos (Unesco).

Pasak F.X.Micheloud (1997) eilučių ir stulpelių taškų palyginimą galima įgyvendinti kampo (ašių susikirtimo atžvilgiu) pagalba:

- jeigu kampas smailus ( $<90^\circ$ ), dvi lyginamos charakteristikos yra koreliuotos;
- jei kampas bukas, taškai yra koreliuoti neigiamai;
- jei kampas status, tarp taškų nėra statistiškai reikšmingo ryšio.

Metodo trūkumai:

- Didesnio mato (m-kintamųjų) lentelių atitikties analizės dimensijos nėra lengvai interpretuojamos kaip kitų tipų faktorinės analizės.
- Tai yra neparametrinis metodas, t.y. jo rezultatų reikšmingumas nėra toks didelis kaip parametrinių metodų.

- Jei tyrimas atliekamas su tolydžiais duomenimis, jie turi būti suskirstyti į intervalus (kategorijas). Tai gali lemti informacijos praradimą. Skirtingas tolydžių duomenų intervalų pasirinkimas taip pat gali daryti įtaką vėlesnių rezultatų interpretavimui. Dėl šių priežasčių dažniau yra pasirenkami kiti metodai labiau tinkantys tolydiems duomenims.

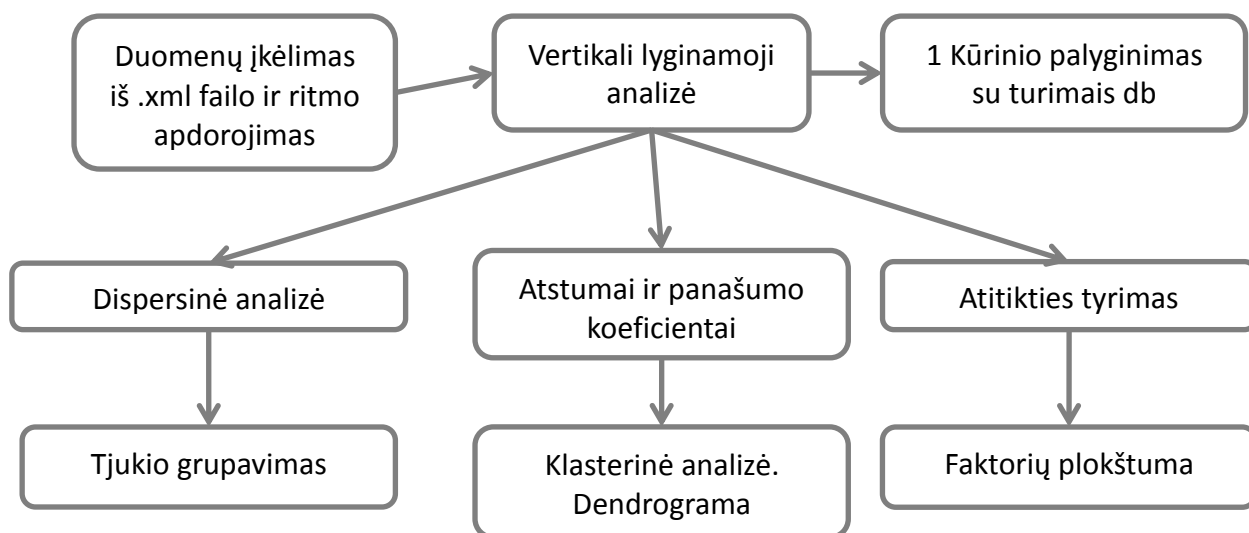
Metodo privalumai:

- Atitikties tyrimas yra neparametrinis metodas, nereikalaujantis pasiskirstymo prielaidų, priešingai nei faktorinė analizė.
- Jis dažniausiai naudojama diskretiems kintamiesiems, kurie turi daug kategorijų. Lentelėms su 2-3 kategorijomis metodas nėra toks informatyvus. Tokiu atveju geriau taikyti kitus metodus.
- Atitikties analizėje teigiama, kad chi-kvadrato atstumai tarp lyginamų dydžių gali būti traktuojami kaip koreliacija tarp kintamųjų.

### 1.3.2 TYRIMO ĮGYVENDINIMAS SAS APLINKOJE

Programinė įranga konkretizuojama vertikaliam duomenų pjūviui, įgyvendinami metodai nereikalaujantys didelio vartotojo įsikišimo.

Pradiniame etape įkeliamas duomenų failas, iš kurio duomenys paruošiami tyrimui, paliekama informacija tik apie ritmo savybes, sukuriamos ritmo chromosomos. Duomenys įkeliami xml žemėlapiu pagalba (angl. *xml map*), jeigu yra tam tikrų netikslumų, apdorojimo ir paruošimo procese gali dalyvauti vartotojas. Naujoje formoje į atskiras skiltis sudedami metodai pagal jų reikšmių pobūdį. Programos struktūra atrodo taip:



**Pav. 1.1 Programos struktūra**



Įkėlus duomenis apdorojimui, pateikiamos aprašomosios charakteristikos ir tikimybinis pasiskirstymas vertikaliam duomenų pjūviui. Remiantis gautais tikimybiniais pasiskirstymais, įgyvendinama parametrinė dispersinė analizė, kurios vizualizacija – Tjukio grupavimas. Toje pačioje programos skiltyje pagal tam tikrą parametą įgyvendinama ir neparametrinė dispersinė analizė. Antroje programos skiltyje apskaičiuojami turimų duomenų vektorių atstumai ir panašumo koeficientai, kurių pagalba gaunama klasterinės analizės vizualizacija – dendrograma. Trečiojoje skiltyje pasirinkimui pateikiamas paprastas atitikties tyrimas ir daugialypis atitikties tyrimas. Šių tyrimų vizualizacija yra pagrindinis interpretavimo objektas. Norint į palyginamąją analizę įtraukti papildomą kūrinių ar kūrinių grupę, reiktų jo dažnių vektorių įtraukti į visų duomenų *xls* failą.

## 1.4 TYRIMO TIKSLAI

Muzikinių kūrinių ritmai paverčiami duomenimis kuriuos galima analizuoti statistiniais metodais. Pasirinktais duomenų pjūviais tiriamos ritmų savybės, atliekamas muzikinių kūrinių grupių palyginimas, jų apjungimas į didesnius darinius, vertinamas grupių panašumas, tikrinamas metodų tikslumas lyginant kūrinių su grupe. Tiriama Čiurlionio ankstyvoji, vėlyvoji kūryba, jo fortepijoninė muzika, papildomai lyginama su lenkų ir dzūkų liaudies muzika. Šioms kūrinių grupėms tirti taikomi metodai:

- Dispersinė analizė;
- Panašumo įvertinimas koeficientų ir atstumų pagalba;
- Klasterinė analizė;
- Paprastoji ir daugialypė atitikties analizė.

Lygindami išvardintais metodais gautus rezultatus galime išskirti metodus kurie suteikia daugiausiai informacijos ir yra tinkami tyrimams muzikoje ar bent jau pasirinktoje kūrinių grupėje.

Jei metodų rezultatai būtų informatyvūs, tai reikštų priartėjimą prie gebėjimo statistinių metodų pagalba atskirti stilių, muzikinę kryptį ar nustatyti kūrinių žanrą. Palyginamieji metodai galėtų dalyvauti plagiatų tyrime. Taip pat šie metodai galėtų padėti spręsti muzikinių kūrinių trūkstančių dalių atkūrimo problemą.

## 2 TRIAMOJI DALIS

### 2.1 DUOMENŲ STRUKTŪRA IR ANALIZUOJAMI PJŪVIAI

Xml žemėlapiu pagalba (žr. priedą nr.2) nuskaitomi duomenys, reikalingi ritmui apibendrinti ir atmesti tam tikras specifines natų grupes:

**Lentelė 2.1 Naudojamų muzikinių parametrų sąrašas**

Nr.	Parametras	Igyjamos reikšmės arba paaiškinimas
1	Taktas	1,2,3,...,n
2	Metras	2/4, 4/4, 3/4, 6/8
3	Tipas	Pusinė, ketvirtinė, aštuntinė,...
4	Taškas	1, 2 (natos prailginimas pusę, 0.75 jos vertės)
5	Ryšys	Besitęsianti nata, ritme tolygi pauzei
6	Nata	C,D,E,F,G,A,B (natos nebuvimas reiškia pauzę).
7	Oktava	2,3,4,5
8	Alteracija	1, -1, 2,-2
9	Akordas	Natos skambančios vienu metu ritmui reikšmės neturi
10	Trumpoji nata	Muzikinis papuošimas, neužimantis vietos ritme
11	Faktinė nata	Realus natos ilgis užimamas takte (specifinių natų grupėse)
12	Normalioji nata	Tikrasis natos ilgis (specifinių natų grupėse)
13	Apjungimas	Besitęsianti nata (jei pakartojama), ritme tolygi pauzei

Pagrindiniai muzikiniai ritmo parametrai: taktas, metras, tipas, taškas, nata. Kiti parametrai naudojami netinkamų duomenų atmetimui (žr. lentelę 2.1). Duomenų paruošimo procese pastebėta, kad muzikiniuose .xml tipo failuose pasitaiko klaidų, dėl to gali būti reikalingas vartotojo įsikišimas į duomenų apdorojimo procesą.

Iš duomenų atmetami taktai su specifiniais ritmais, akordų dalys ir muzikiniai ornamentai. Natos turinčios apjungimą arba ryšį paverčiamos pauzėmis. Pasirinktinai atmetamos šešiasdešimt ketvirtinės natos, nes jų pasitaiko labai retai ir tušti taktai, nes šiuo atveju jie nėra reikšmingi. Pateikiami tolimesni duomenų tvarkymo žingsniai:

- Duomenys tvarkomi trisdešimt antrinės natos tikslumu, t.y. kiekvienam natos ilgiui priskiriama skaitinė vertė, 32-inė nata atitinka 1, 16-inė atitinka 2, aštuntinė 4 ir t.t.
- Jeigu nata su vienu tašku, ji prailginama pusę savo vertės, jei su dviem – 0.75 karto.

- Priklausomai nuo metro, sudaromos atitinkamo ilgio chromosomos, pavyzdžiui, 3/4 metre taktas sudarytas iš 24 elementų.
- Kadangi muzikiniai kuriniai gali turėti daugiabalsiškumą, t.y. turi kelias melodines linijas, vienas taktas gali būti padalintas į 2 ar 3 skirtingo ritmo taktus.
- Jei yra neatitikimų muzikiniame *xml* faile informacija apie klaidingus taktus patalpinama į atskirą *sas* failą. Likusieji teisingi taktai užrašomi binarine forma: 1 – reiškia natos buvimą arba pradžia, 0 – natos tęsinį arba pauzę.
- Jei norima pakeisti takto tikslumą į 16-inės natos, automatiškai atmetamos 32-inės natos, nes iš takto chromosomos atmetamas kas antras elementas.

Pavyzdžiui paveiksle 2.1 pateikiama skirtingų melodijos atžvilgiu, tačiau vienodo ritmo, taktų kodavimas 3/4 metre 16-inės natos tikslumu. Šiuo atveju chromosomos ilgis lygus 12.



**Pav. 2.1 Muzikinio ritmo taktų kodavimas**

Suskaičiuojami chromosomų dažniai visame kūrinyje ir gaunama kūrinį ar kūrinijų grupę apibūdinanti formulė – taip gaunamas **horizontalus duomenų pjūvis**.

Chromosomų įvairovė yra didelė, dėl to horizontalaus pjūvio duomenis lyginti yra sudėtingiau, ir kai kurių metodų atveju didelė dalis duomenų turi būti atmesta. Dėl šios priežasties naudojamas **vertikalus duomenų pjūvis**, kai analizuojamas kiekvienas takto kirtis. Analogiškai R. Kašponio (1992) tyrimuose analizuojama vadinamoji pulsacija. Tokiu atveju turime  $n$  kintamųjų, kai  $n$  – yra chromosomos ilgis, ir kiekvienas kintamasis atitinka konkretaus  $i$ -tojo takto kirčio dažnį. Šis dažnis yra lygus visų taktų, kuriuose  $i$ -tajame kirtyje pasirodo nata (chromosomoje - 1), dažnių sumai. Gaunamas  $n$  elementų ilgio vektorius, kuris laikomas kūrinio ar kūrinijų grupės charakteristika, arba ritmo formule.

Šiam tyrimui naudojamos tokios kūrinijų grupės:

- dzūkų liaudies muzikos rinkiniai,
- lenkų liaudies muzikos rinkiniai,
- Čiurlionio fortepijoninės muzikos rinkiniai,
- Čiurlionio ankstyvosios kūrybos rinkiniai,
- Čiurlionio vėlyvosios kūrybos rinkiniai.

Čiurlionio fortepijoninės, ankstyvosios ir vėlyvosios kūrybos ritmai atrinkti iš knygos „Kūriniai fortepijonui“ (2004). Duomenų grupės, vadinamos „fortepijonine“ kūryba, kūriniai

atrinkti atsitiktinai, tačiau pastebėta, kad visi jie priklauso vėlyvosios Čiurlionio kūrybos etapui. Todėl fortepijoninė ir vėlyvoji kūryba, liaudies kūryba naudojama lyginamųjų metodų rezultatų vertinimui. Dzūkų liaudies muzikos rinkiniai – iš knygų „Druskininkų dainos“ (1972), „Dzūkų melodijos“ (1981).

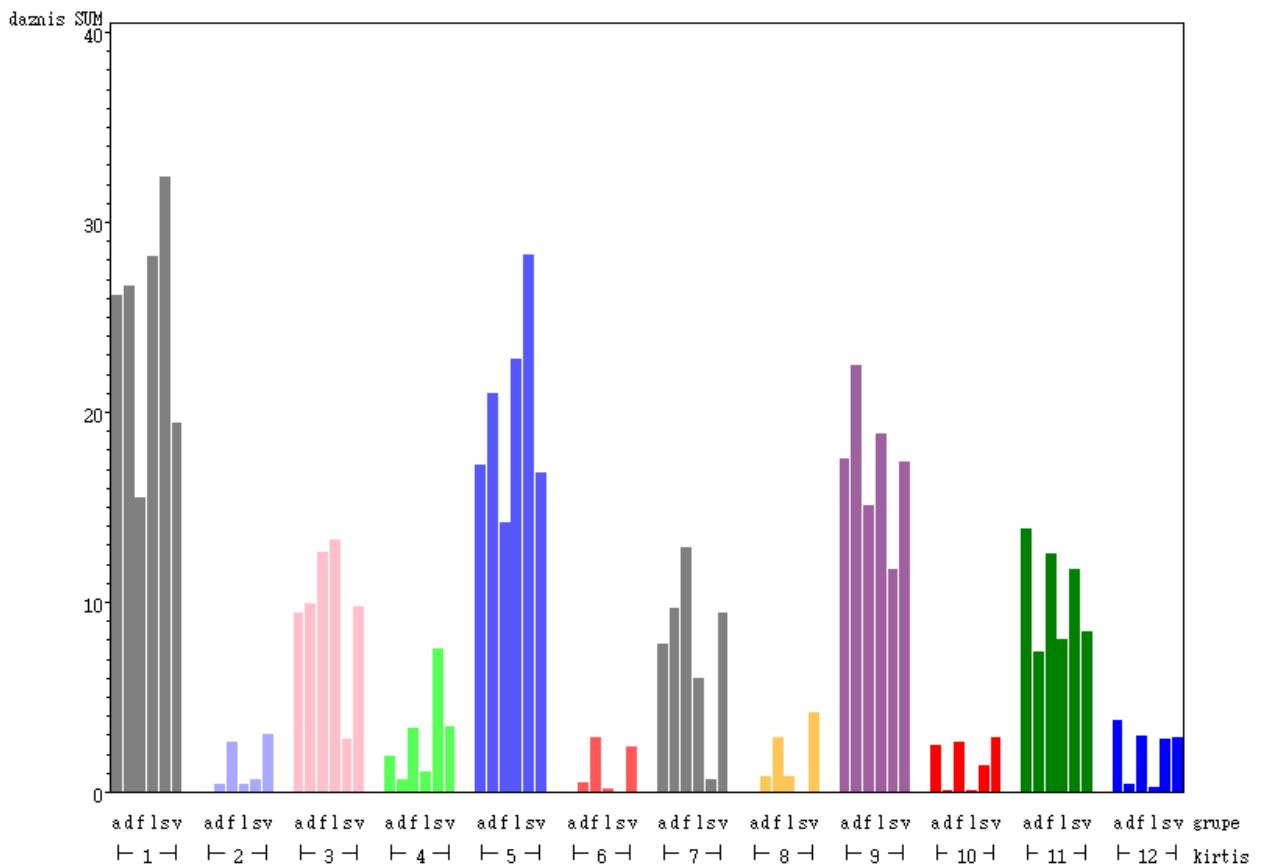
Atskirai ankstesniuose tyrimuose buvo analizuota Čiurlionio simfoninė poema „Jūra“ ir visų jos instrumentų partijos. Dalis duomenų atrinkta ir paruošta rankiniu būdu, dalis įkelta iš muzikinių *xml* failų. Toliau pateikiamų modelių pagalba įgyvendinama šių kūrinių grupių ritmo formulių lyginamoji analizė.

## **2.2 MUZIKINIO RITMO STATISTINĖS ANALIZĖS MODELIAI**

### **2.2.1 APRAŠOMOSIOS CHARAKTERISTIKOS**

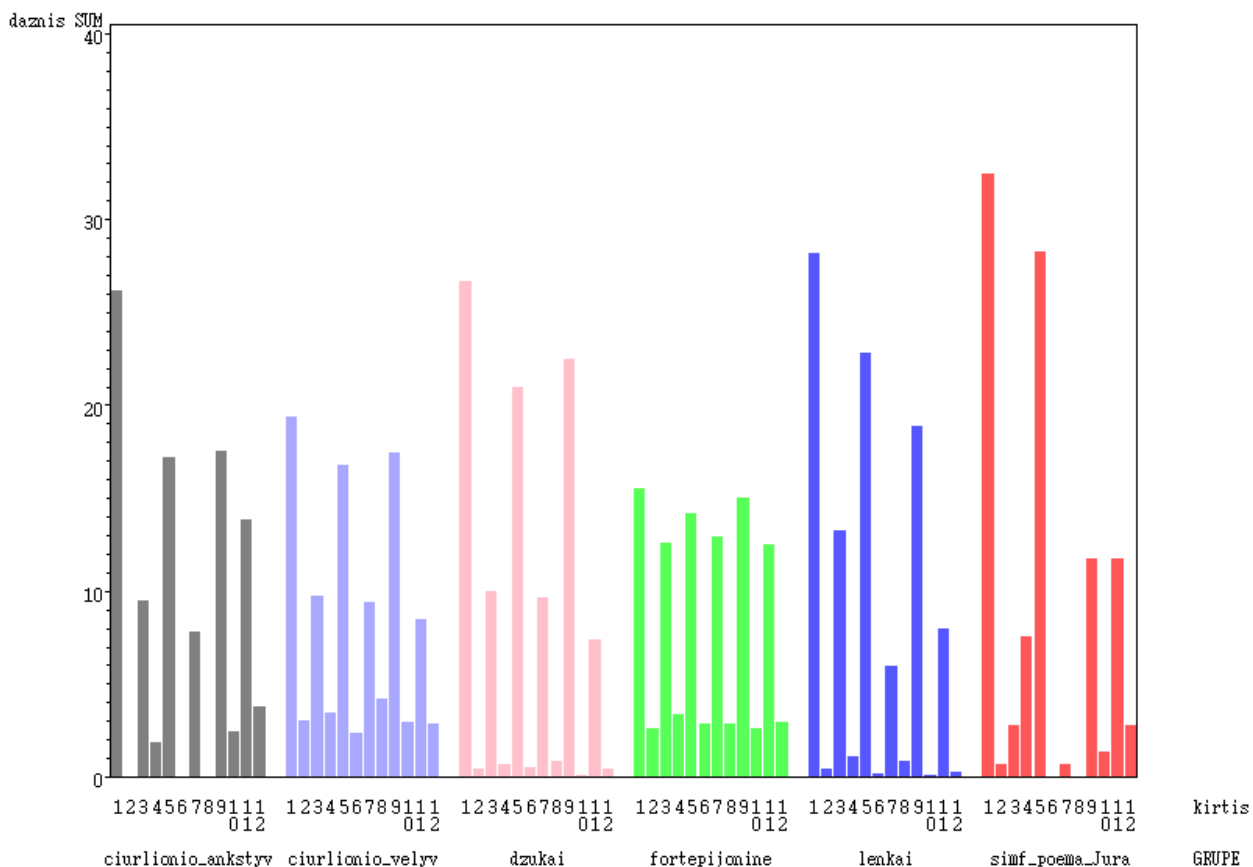
Palyginamajai analizei naudojami normuoti duomenys, t.y. pulsacijos vektoriai parodo kiekvienos takto dalies stiprumą takte procentais, todėl vektoriaus elementų suma lygi 100.

Procedūra *gchart* sukuriama grafikais, kuriuose sugrupuoti visų kūrinių ir kūrinių grupių taktų kirčių pasiskirstymai. Pirmame grafike lyginamos duomenų grupės kiekvieno takto kirčio atžvilgiu (žr. Pav. 2.2). Duomenų grupės surašytos sutrumpinimais: a – ankstyvoji M. K. Čiurlionio kūryba, d – dzūkų liaudies muzika, f – fortepijoninė Čiurlionio kūryba, l – lenkų liaudies muzika, s – Čiurlionio simfoninė poema „Jūra“, v – vėlyvoji Čiurlionio kūryba. Jei lyginsime pagal stipriąsias trijų ketvirtinių metro takto dalis (1, 5, 7), kurios kaip matome turi didžiausius dažnius, simfoninė poema „Jūra“ turi stipriausią pulsaciją 1 ir 3 takto dalyse, o fortepijoninė muzika – silpniausią pulsaciją, lyginant su tiriamomis kūrinių grupėmis. 2, 4, 6, 8, 10, 12 takto dalyse dzūkų ir lenkų liaudies muzika turi mažiausius dažnius.



**Pav. 2.2 Taktų kirčių pasiskirstymas pagal grupes**

Jeigu norime lyginti skirstinio tankio kreives formas atžvilgiu, galime interpretuoti antrąją grafiką, kuriame visi takto kirčių dažniai pateikti kiekvienai kūrinio grupei (žr. Pav. 2.3). Lyginant atskirai stipriųjų ir silpnųjų takto dalių atžvilgiu, simfoninės poemos „Jūra“, ankstyvosios Čiurlionio kūrybos ir lenkų liaudies muzikos pulsacija netolygi. Lyginant atskirai stipriąsias takto dalis dzūkų liaudies muzika ir vėlyvoji Čiurlionio kūryba turi artimas histogramas formas atžvilgiu.



**Pav. 2.3 Taktų kirčių pasiskirstymas kiekvienoje kūrinių grupėje**

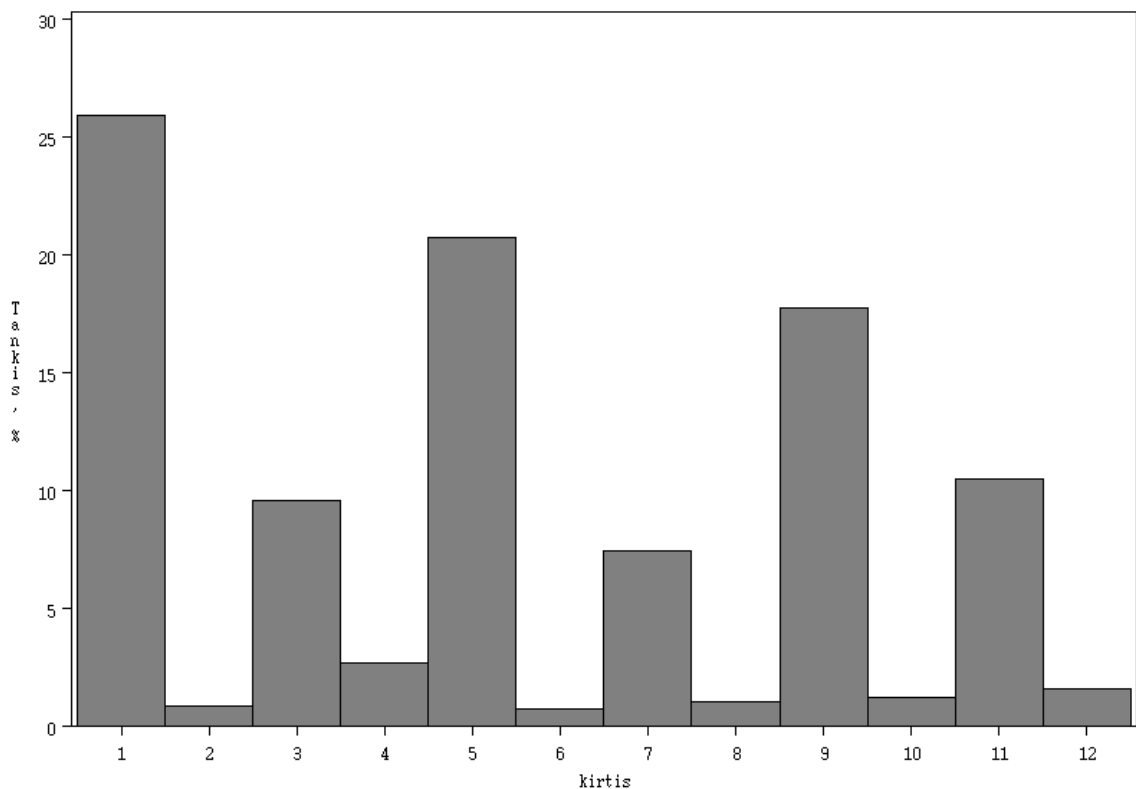
Pasiskirstymo tyrime tikrintos suderinamumo hipotezės, tačiau visos jos buvo atmestos, o tai galima nuspėti ir iš turimos histogramos, kuri kartu su aprašomosiomis charakteristikomis pateikiama SAS Base procedūros *univariate* pagalba (žr.Pav. 2.4). Aprašomosios charakteristikos apskaičiuojamos visiems duomenims kartu. Pavyzdžiui lentelėje 2.2 pateikti kvantiliai rodo, kad 25% pasiskirstymo sudaro pirmasis takto kirtis, o 50% - takto kirčiai iki 5-ojo.

**Lentelė 2.2 Taktų kirčių kvantiliai**

Kvantiliai	
Kvantiliai	Įverčiai
100% Max	12
99%	12
95%	11
90%	11
75% Q3	9
50% Mediana	5

Kvantiliai	
Kvantiliai	Įverčiai
25% Q1	1
10%	1
5%	1
1%	1
0% Min	1

Pateiktoje histogramoje kaip ir anksčiau aptartuose grafikuose matome, kad didžiausius dažnius turi muzikiniu aspektu stipriosios takto dalys: 1, 5, 9.



**Pav. 2.4 Taktų kirčių histograma**

## 2.2.2 NEPRIKLAUSOMUMO TYRIMO MODELIS

Nepriklausomumui tirti taikoma parametrinė dispersinė analizė ir neparametriniai kriterijai. Procedūros *glm* pagalba įgyvendinamas kūrinių grupių nepriklausomumo tyrimas,

pateikiamas Tjukio grupavimas. Šie metodai taikomi ir normuotiems ir nenormuotiems duomenų vektoriams. Pavyzdžiui pateikti nenormuotų duomenų hipotezės apie vidurkių lygybę tarp kūrinių grupių tikrinimo rezultatai (žr. lentelę 2.3). Su pasirinktu reikšmingumo lygmeniu 0,05, hipotezė atmesta su p-reikšme pateikta lentelėje. Tačiau remiantis R-kvadrato reikšme gauti rezultatai paaiškina tik 1,4% variacijos. Žemiau lentelėje 2.4 pateikiamas Tjukio grupavimas. Matome du galimus grupavimo variantus:

- Čiurlionio kūrybos grupė ir liaudies kūrybos su simf. poema „Jūra“ grupė (A);
- Viena grupė sudaryta iš Čiurlionio ankstyvosios, vėlyvosios kūrybos ir dzūkų liaudies muzikos (B).

**Lentelė 2.3 Dispersinės analizės rezultatai**

Nuokrypų šaltinis	LL	Nuokrypių kvadratų sumos	Nuokr. kvadratų vidurkiai	Fišerio statistika	p-reikšmė
Faktorius <i>grupe</i>	5	408,2154	81,64308	6,6	<,0001
Atsitiktinis faktorius	2216	27394,87371	12,36231		
Visi faktoriai	2221	27803,08911			

**Lentelė 2.4 Tjukio grupavimo lentelė**

Tjukio grupavimas	Vidurkis	N	Grupe
A	6,0692	477	fortepijonine
A			
B	5,8035	341	ciurlionio_ankstytv
B			
B	5,7943	423	ciurlionio_velyv
B			
B	5,3035	369	dzukai
C			
C	4,9641	334	lenkai
C			
C	4,9640	278	simf_poema_Jura



Kiti parametrinės analizės rezultatai pateikti priede Nr.6. Patikrinus dispersinės analizės prielaidas, nustatyta kad modelio rezultatai nėra patikimi ir paaiškinta tik labai mažą sklaidos dalį.

Procedūra *npar1way* pateikia pasirinktus koeficientus (angl. *scores*) ir galutinius taikomų kriterijų rezultatus - chi-kvadrato statistiką ir jos įvertinimą. Kaip ir prieš tai pateiktame pavyzdyje tiriama hipotezė apie vidurkių lygybę (šiuo atveju keturiais kriterijais). Wilcoxon koeficientams taikomas Kruskal-Wallis kriterijus, koeficientai yra rangų sumos (žr.lentelę 2.5).

Visų taikytų kriterijų atveju hipotezė atmesta, tai reiškia kad taktų kirčių skirtumas tarp tiriamų kūrinių grupių kategorijų yra reikšmingas (žr.lentelę 2.6).

**Lentelė 2.5 Wilcoxon rangų sumos**

Wilcoxon koeficientai (rangų sumos) kintamajam <i>kirtis</i> , sugrupuotam pagal kintamąjį <i>grupe</i>					
Grupė	N	Rangų suma	Tikėtina rangų suma pagal $H_0$	Stand. nuokrypis pagal $H_0$	Rangų vidurkis
ciurlionio_ankstyv	341	393176,00	379021,50	10739,7099	1153,00880
ciurlionio_velyv	423	489744,00	470164,50	11697,8683	1157,78723
dzukai	369	394479,50	410143,50	11088,4759	1069,05014
fortepijonine	477	576287,50	530185,50	12234,2646	1208,14990
lenkai	334	336886,50	371241,00	10648,6657	1008,64222
simf_poema_Jura	278	279179,50	308997,00	9858,0608	1004,24281

Medianos koeficientų rezultatai, kurie paremti reikšmėmis didesnėmis už medianą, nepriklausomumo hipotezę taip pat atmeta. Nors neparametriniai metodai nereikalauja prielaidų susijusių su kintamojo pasiskirstymu, jų rezultatai gali būti efektyvesni tam tikriems skirstiniams: Medianos koef. rezultatai ypač efektyvūs – simetriniams skirstiniams ir su „sunkiomis uodegomis“, Van der Waerden koef. (sutrump. VW) rezultatai – normaliojo skirstinio atveju, „Savage“ koeficientų rezultatai – eksponentinio skirstinio atveju (Scores for Linear Rank and One-Way ANOVA Tests ). Lentelės 2.6 rezultatai rodo, kad su reikšmingumo lygiu 0,05, nepriklausomumo hipotezė atmesta visų taikytų kriterijų atžvilgiu.

**Lentelė 2.6 Neparametrinių kriterijų rezultatų palyginimas**

Kriterijus	Chi-kvadrato statistika	p-reiškė
Wilcoxon	33,4021	<,0001
Medianos	41,8246	<,0001
VW	36,5181	<,0001
„Savage“	35,2390	<,0001

### 2.2.3 PANAŠUMO KOEFICIENTAI IR KLASTERINĖ ANALIZĖ

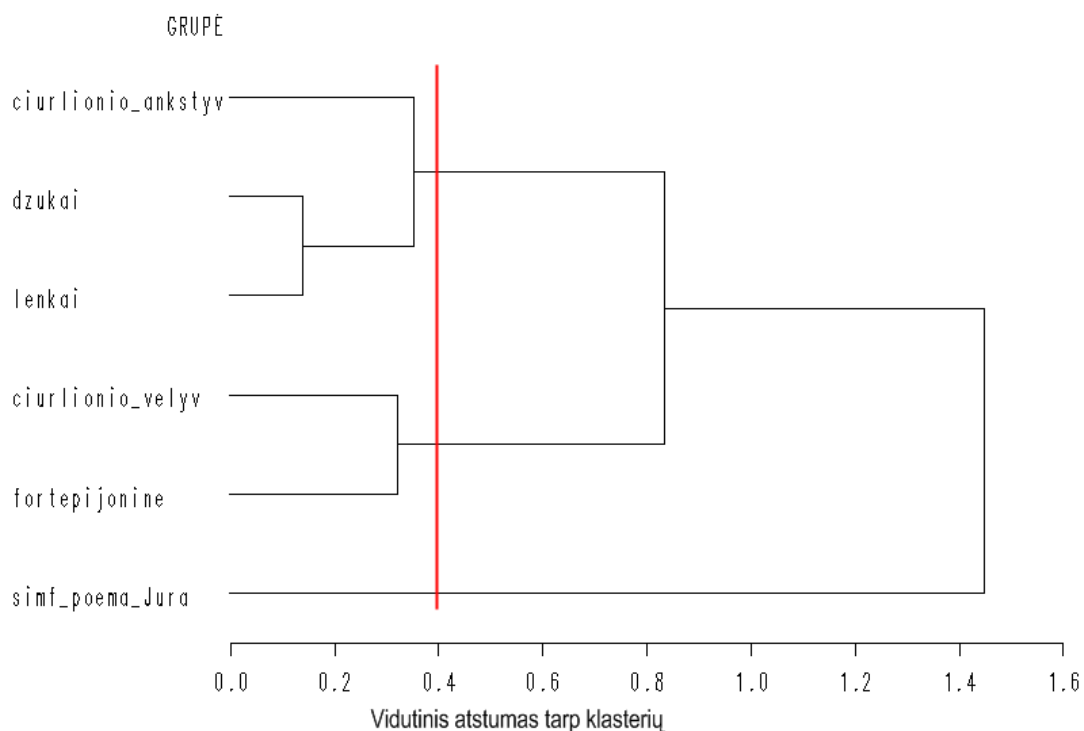
Panašumo koeficientai naudojami turimų kūrinių grupių panašumo lygiui įvertinti. Pateikiami Euklido ir formos atstumai, panašumo ir koreliacijos koeficientai palyginimui kaip skirtingų tipų dydžiai, bei pasirinkti kaip koeficientai duodantys tiksliausius rezultatus muzikiniu aspektu.

**Lentelė 2.7 Formos koeficiento reikšmės**

Nr.	Kūrinys ar kūrinių grupė	<i>ciurlionio_ankstyv</i>	<i>ciurlionio_velyv</i>	<i>dzukai</i>	<i>fortepijonine</i>	<i>lenkai</i>	<i>simf_poema_Jura</i>
1	<i>ciurlionio_ankstyv</i>	0,0000	.	.	.	.	.
2	<i>ciurlionio_velyv</i>	10,6547	0,0000	.	.	.	.
3	<i>dzukai</i>	10,1995	11,8182	0,0000	.	.	.
4	<i>fortepijonine</i>	13,8393	8,1827	17,4217	0,0000	.	.
5	<i>lenkai</i>	10,4238	13,5209	6,6304	18,7901	0,0000	.
6	<i>simf_poema_Jura</i>	18,1620	22,7546	20,2307	27,9530	17,3868	0

Pagal lentelėje 2.7 pateiktas formos koeficiento reikšmes galime teigti, kad mažiausias atstumas tarp vertikalaus pjūvio duomenų vektorių įgyjamas lyginant lenkų ir dzūkų liaudies muziką, Čiurlionio vėlyvą ir jo fortepijoninę muziką. Simfoninė poema „Jūra“ labiausiai nutolusi lyginant su kitų grupių koeficientais, nuo Čiurlionio vėlyvosios, fortepijoninės muzikos ir dzūkų liaudies muzikos. Pagal analizuojamus atstumus taip pat pateikiami klasterinės analizės rezultatai – grupavimo medis arba dendrograma. Taikant klasterinės analizės vidurkių metodą dendrograma pateikia analogišką grupavimą metodais *disratio*, *Euclid*, *shape*, *nonmetric*, *overlap*, jei pasirenkame dviejų elementų klasterio dydį. Tačiau koreliacijos koeficiento paversto

Euklido atstumu atveju, grupavimas skiriasi nuo gauto prieš tai minėtais metodais. Pavyzdžiui pateiktas kūrinių grupių klasterizavimas, naudojant panašumo koeficientą paverstą atstumu, t.y.  $simratio \rightarrow disratio$  (žr.Pav. 2.5). Matome, kad gaunami du dviejų elementų dydžio klasteriai: viename – dzūkų ir lenkų liaudies muzika, kitame Čiurlionio vėlyvoji ir fortepijoninė muzika.

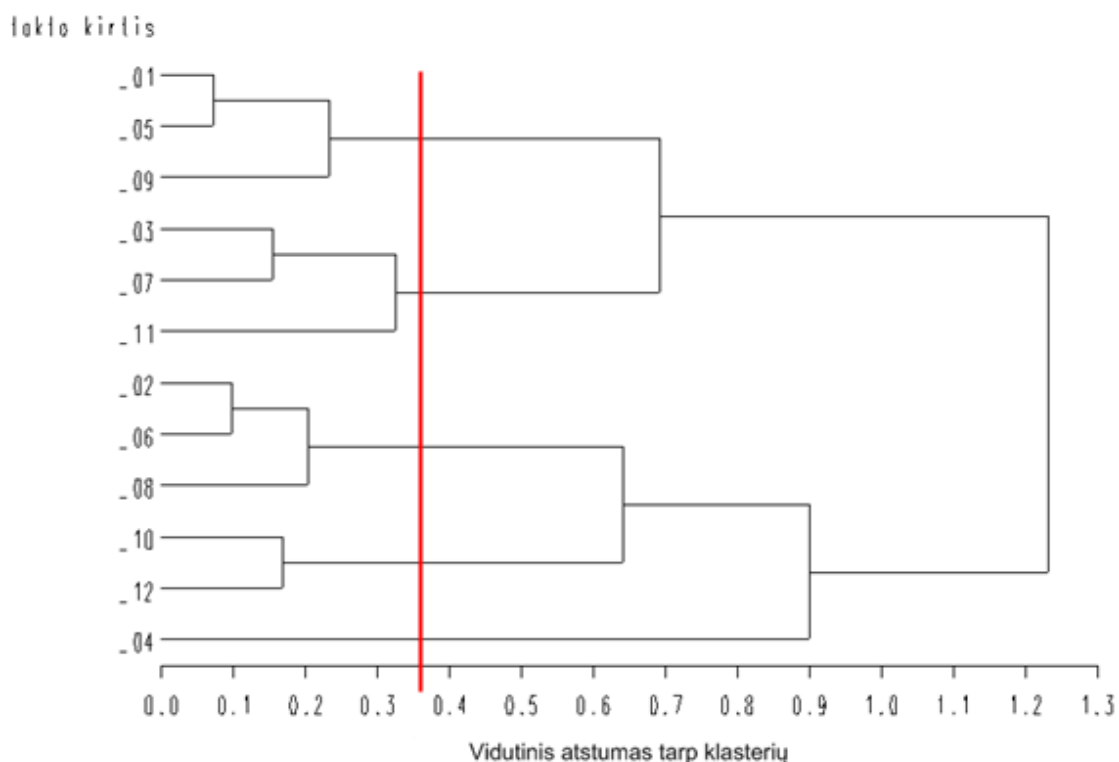


**Pav. 2.5 Kūrinių grupių dendrograma pagal panašumo koeficientą**

Jei pasirinktume klasterius kirtimo linijos pagalba, pavyzdžiui kaip pažymėta Pav. 2.5, turėtume tris klasterius:

- Čiurlionio ankstyvoji, dzūkų liaudies ir lenkų liaudies muzika;
- Čiurlionio vėlyvoji ir fortepijoninė muzika;
- Simfoninė poema „Jūra“.

Tam kad įvertintume takto kirčių svarbumą, galime interpretuoti jų grupavimą, kuris taip pat atliekamas naudojant pulsacijos stiprumo vektorius. Tik šiuo atveju kūrinių grupių vektoriai transponuojami ir kiekviena grupė atitinka atskirą stebėjimą tiriant takto kirčių pulsaciją. Jei kirtimo liniją pasirinksim kaip parodyta žemiau (žr. Pav. 2.6), turėsime penkis klasterius. Pirmasis susideda iš 1, 5 ir 9 kirčių, antrasis – iš 3, 7 ir 11 kirčių, trečiasis yra pirmasis pastumtas vienetas – 2, 6, 8 kirčiai. Jei pastumtume kirtimo liniją į dešinę, taip kad iš viso būtų du klasteriai – turėtume stipriųjų takto dalių grupę (ją galėtume gauti sumažindami duomenų vektorius aštuntinės natos tikslumu) ir silpnųjų takto dalių grupę.



**Pav. 2.6 Takto kirčių grupavimas pagal panašumo koeficientą**

Papildomai tyrimo metu buvo naudojamas Kullback-Leibler atstumas dar vadinamas KL-divergencija. Kadangi šis koeficientas reikalauja, kad duomenyse nebūtų trūkstatų reikšmių, naudojami tik stebėtus dažnius turintys taktų kirčiai. Lentelėje 2.8 pateiktos divergencijos reikšmės simfoninės poemos „Jūra“ palyginimui su turimomis kūrybių grupėmis. Pirmajame stulpelyje pateiktos koeficiento reikšmės, apskaičiuotos atmetus 6 ir 8 takto kirčius, todėl Čiurlionio ankstyvoji muzika negali būti lyginama šiuo atveju. Antrajame stulpelyje atmesti 2, 6 ir 8 takto kirčiai, trečiajame – lyginama tik stipresniųjų takto dalių atžvilgiu (1, 3, 5, 7, 9 ir 11). Šis koeficientas yra atstumas apibrėžtas  $[0; \infty)$ , kur 0 reikšmė įgyjama kai pasiskirstymai yra labai panašūs arba identiški. Pirmuoju atveju didžiausią panašumą į simfoninę poemą turi Čiurlionio vėlyvoji muzika, o atmetus antrąjį takto kirtį – ir Čiurlionio ankstyvoji muzika. Lyginant pagal stipresniąsias takto dalis, simfoninei poemai artimiausia lenkų liaudies muzika.

**Lentelė 2.8 KL-divergencija simfoninės poemos „Jūra“ palyginimui**

Kūrybių grupė	Kullback-Leibler divergencija		
	1	2	3
Čiurlionio ankstyvoji muzika	-	0,2752	0,2228
Čiurlionio vėlyvoji muzika	0,3234	0,3034	0,3076
Dzūkų liaudies muzika	0,4929	0,4953	0,2677
Čiurlionio fortepijoninė muzika	0,4592	0,4459	0,4699
Lenkų liaudies muzika	0,4041	0,4059	0,2096

Kadangi šis koeficientas yra nesimetrinis, palyginsime gautus rezultatus su simetrine jo forma – J-divergencija. Šis parametras įgyja reikšmes tame pačiame intervale. Palyginimui pateikiama analogiška rezultatų lentelė (žr lentelę Nr.2.9). Pateikti rezultatai rodo, kad pirmuoju atveju simfoninei poelai artimiausia lenkų liaudies ir Čiurlionio ankstyvoji muzika, antruoju atveju – Čiurlionio ankstyvoji ir vėlyvoji muzika, stipresniųjų takto dalių atžvilgiu – lenkų liaudies muzika ir ankstyvoji Čiurlionio kūryba.

**Lentelė 2.9 Simetrinė J-divergencija simfoninės poemos „Jūra“ palyginimui**

Kūrinių grupė	Jeffreys divergencija		
	1	2	3
Čiurlionio ankstyvoji muzika	-	0,6338	0,5550
Čiurlionio vėlyvoji muzika	0,8288	0,7884	0,8132
Dzūkų liaudies muzika	1,0260	1,0299	0,6832
Čiurlionio fortepijoninė muzika	1,2002	1,1816	1,2535
Lenkų liaudies muzika	0,8216	0,8244	0,5185

## 2.2.4 ATITIKTIES TYRIMO MODELIS

Atitikties analizė įgyvendinama normuotai dažnių lentelei, kur stulpeliuose išdėstyti kintamieji – kūrinių grupės, o eilutėse – taktų kirčių kategorijos (tiek stulpelius tiek eilutes galima vadinti kategoriniais kintamaisiais). Tiriamos 6 kūrinių grupės ir jų pasiskirstymas 12-oje takto kirčių. Atitikties analizė pateikia kategoriniams duomenims pagrindinius faktorius - dimensijas, kur kaip įprasta du pirmieji faktoriai paaiškina didžiąją dalį taškų sklaidos (angl. *cloud*). Pavyzdžiui lentelėje 2.10 pirmoji dimensija paaiškina 55,34% taškų pasiskirstymo, o antroji – 29,53%. Gauname, kad pasirinktos dvi dimensijos paaiškina 84,86% visų taškų sklaidos.

**Lentelė 2.10 Inercija ir chi-kvadrato statistikos atitikties tyrimo dimensijoms**

Inercija ir Chi-kvadrato statistikos					
Singuliari reikšmė	Atstovaujanti inercija	Chi-kvadratas	Proc.	Sukaupta Proc.	11 22 33 44 55 -----+-----+-----+-----+-----
0,28108	0,07900	47,4023	55,34	55,34	*****
0,20531	0,04215	25,2921	29,53	84,86	*****
0,12388	0,01535	9,2074	10,75	95,61	*****
0,06011	0,00361	2,1680	2,53	98,14	*

Inercija ir Chi-kvadrato statistikos					
Singuliari reikšmė	Atstovaujanti inercija	Chi- kvadratas	Proc.	Sukaupta Proc.	11 22 33 44 55 -----+-----+-----+-----+-----+-----
0,05149	0,00265	1,5905	1,86	100,00	*
Total	0,14277	85,6603	100,00		
Degrees of Freedom = 55					

Pavyzdžiui pateikti stulpelių taškų parametrai, eilučių parametrai ir jų interpretavimas pateiktas rezultatų priede nr.6. Pateiktoje lentelėje 2.11 įvertinama taškų kokybė, inercija abiejų ašių atžvilgiu ir stulpelio masės parametras visų stulpelių atžvilgiu. Kokybės parametras parodo kaip gerai pasirinktos dvi ašys apibūdina konkretaus taško koordinates. Konkretaus taško inercija parodo kokią dalį bendrosios inercijos jis užima, nepriklausomai nuo to kiek dimensijų pasirinkta (Creators of Statistica ). Šiuo atveju matome, kad Čiurlionio ankstyvoji muzika nėra kokybiškai apibrėžta šiomis dviem dimensijomis, taip pat sudaro mažiausią dalį visos inercijos (0,862). Stulpelių masė yra vienoda, nes visų stulpelių dažnių sumos yra vienodos. Didžiausią dalį – 33,7% visos inercijos sudaro simfoninė poema „Jūra“, be to jos koordinacių kokybė yra pakankamai didelė pirmų dviejų ašių atžvilgiu (0,9875). Taip pat didelę inerciją ir didelę koordinacių kokybę turi Čiurlionio fortepijoninė ir vėlyvoji kūryba.

**Lentelė 2.11 Aprašomosios statistikos stulpelių taškams**

Grupė	Kokybė	Masė	Inertia
ciurlionio_ankstyv	0,0991	0,1667	0,0862
ciurlionio_velyv	0,8407	0,1667	0,1600
dzukai	0,8707	0,1667	0,1027
fortepijonine	0,9354	0,1667	0,2080
lenkai	0,8365	0,1667	0,1058
simf_poema_Jura	0,9875	0,1667	0,3372

**Lentelė 2.12 Dalinė inercija stulpelių taškams**

Grupė	Dim1	Dim2
ciurlionio_ankstyv	0,0139	0,0029
ciurlionio_velyv	0,2026	0,0759

Grupė	Dim1	Dim2
dzukai	0,0053	0,2931
fortepijonine	0,3280	0,0444
lenkai	0,0355	0,2333
simf_poema_Jura	0,4148	0,3504

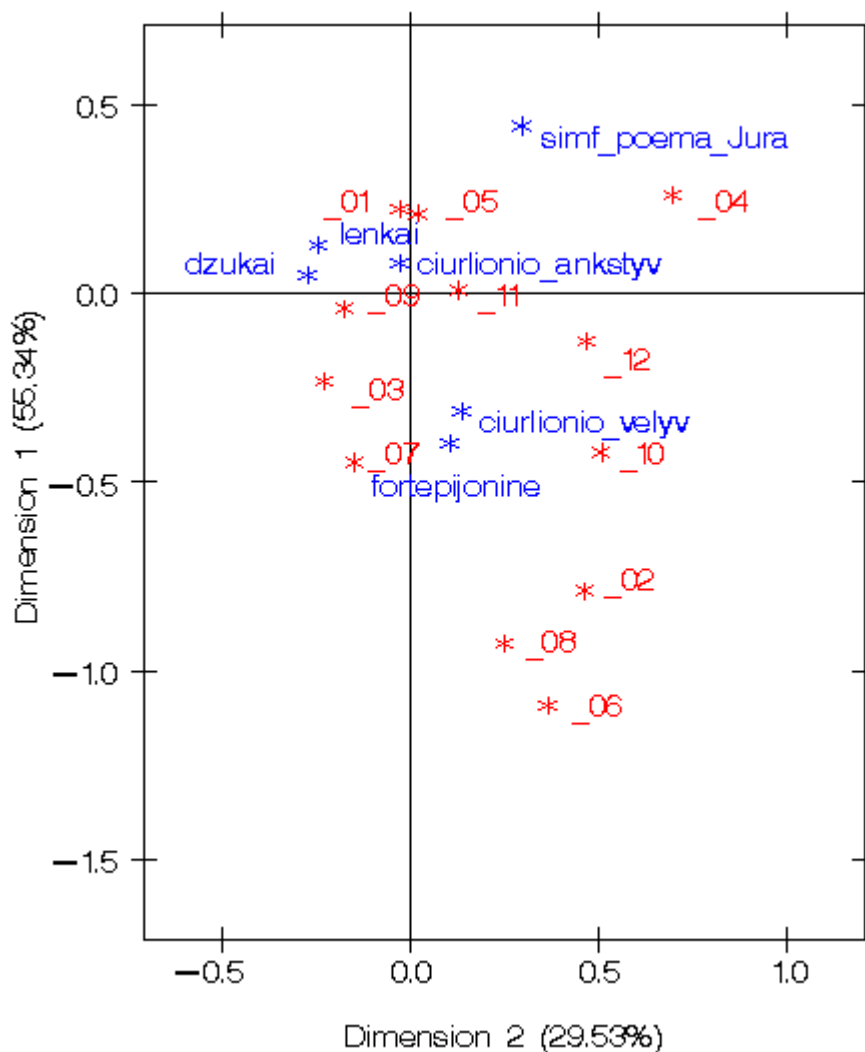
Vertinant koordinacių patikimumą kiekvienos dimensijos atžvilgiu atskirai – didžiąją dalį inercijos ir pirmos ir antros dimensijos sudaro simfoninė poema (žr. lentelę 2.12). Čiurlionio vėlyvoji ir fortepijoninė muzika sudaro didžiausią dalį inercijos pirmojoje dimensijoje, dzūkų ir lenkų liaudies muzika geriausiai apibūdinamos antrosios dimensijos. Nors Čiurlionio ankstyvoji kūryba sudaro mažą dalį visos inercijos ir kiekvienos dimensijos inercijos atskirai, indeksų lentelė, kurią pateikia SAS šios procedūros rezultatuose, rodo, kad geriau šį tašką apibūdina pirmoji dimensija. Taip pat rezultatuose pateikiama kosinusų kvadratų lentelė kiekvienai dimensijai, kur šis parametras dar vadinamas kokybe kiekvienai dimensijai, arba kvadratine „koreliacija“ su kiekviena dimensija (Creators of Statistica ). Pavyzdžiui, lentelėje 2.13 matome, kad simfoninė poema „Jūra“, Čiurlionio vėlyvoji ir fortepijoninė muzika yra labiausiai koreliuotos su pirmąja dimensija, o lenkų ir dzūkų liaudies muzika – su antrąja. Šiuo atveju gauti rezultatai analogiški gautiems lyginant inercijos proporcijas, tačiau taškas gali sudaryti didelę dalį inercijos atitinkamoje dimensijoje, bet jo kokybė tos dimensijos atžvilgiu būti maža.

**Lentelė 2.13 Kosinusų kvadratai stulpelių taškams**

Grupė	Dim1	Dim2
ciurlionio_ankstyv	0,0892	0,0099
ciurlionio_velyv	0,7007	0,1400
dzukai	0,0283	0,8424
fortepijonine	0,8724	0,0630
lenkai	0,1856	0,6509
simf_poema_Jura	0,6807	0,3068

Kadangi atitikties žemėlapyje atstumai tarp taškų gali būti interpretuojami tik kintamojo viduje, palyginsime stulpelių taškus. Kaip ir iš prieš tai pateiktų lentelių, galime pastebėti, kad artimiausi yra Čiurlionio vėlyvoji ir fortepijoninė muzika, bei lenkų ir dzūkų liaudies muzika (žr.Pav. 2.7). Jei vertinsime kampą tarp eilučių ir stulpelių taškų, dzūkų liaudies ritmo pulsacijai

neturi įtakos 7-asis takto kirtis, nes kampas tarp šių taškų dimensijų centro atžvilgiu yra status. Tuo tarpu takto kirčius galėtume sugrupuoti taip: 1 su 5, 3 su 7, ir matoma atskira takto kirčių grupė stipriai nutolusi pirmosios dimensijos atžvilgiu – 2, 6, 8.



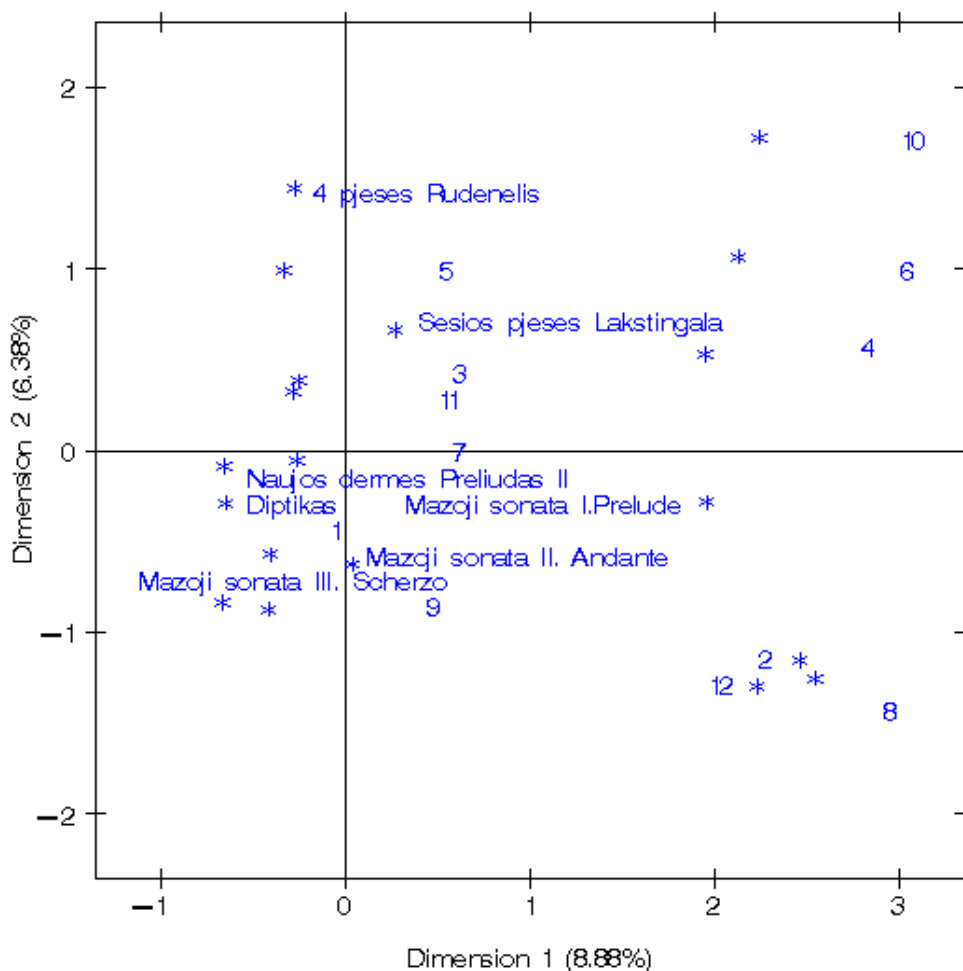
**Pav. 2.7 Atitikties žemėlapis**

Atitiktį analizės prielaida apie stulpelių homogeniškumą kategorijų (eilučių) atžvilgiu priimta, remiantis *Levene* kriterijumi.

Daugialypio atitikties tyrimo atveju, sukuriama *Burt* lentelė, kurioje perskaičiuojami abiejų kategorinių kintamųjų dažniai, ir tolimesni skaičiavimai atliekami analogiškai kaip įprastosios atitikties analizės atveju. Kadangi ši dažnių lentelė simetrinė, pateikiami skaičiavimai tik stulpelių koordinatėms. Šiuo atveju visi taškai gali būti lyginami atstumo atžvilgiu, galime sakyti, kad dvi dimensijos atskiria keturias kintamųjų grupes. Pateiktas fortepijoninės muzikos kūrinių daugialypės atitikties analizės žemėlapis rodo, kad pirmosios dvi dimensijos paaiškina 15,26% inercijos (žr.Pav. 2.8). Pirmajame ketvirtyje yra pjesė „Rudenėlis“ ir 3, 5, 11 takto kirčiai, antrajame ketvirtyje viena iš šešių pjesių „Lakštingala“ ir 4, 6, 11 takto kirčiai,



ketvirtajame – sugrupuoti kūriniai Naujose dermės preliudas II, Diptikas ir Mažoji sonata III Sherzo.

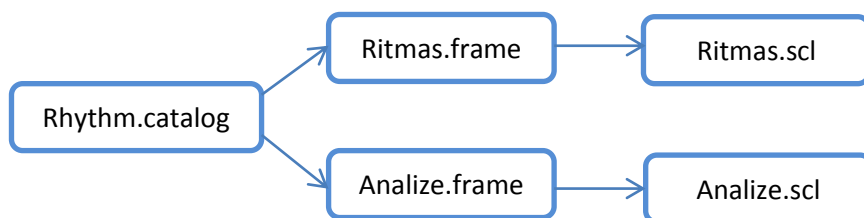


**Pav. 2.8 Daugialypės atitikties analizės žemėlapis**

Šiuo atveju didžiąją dalį pirmosios dimensijos inercijos sudaro pjesė „Rudenėlis“, o antrosios dimensijos – Mažoji sonata I Preliudas. Analogiškai šių kategorijų kokybė yra žymiai didesnė nei kitų tiek bendruoju atveju, tiek dimensijų atžvilgiu.

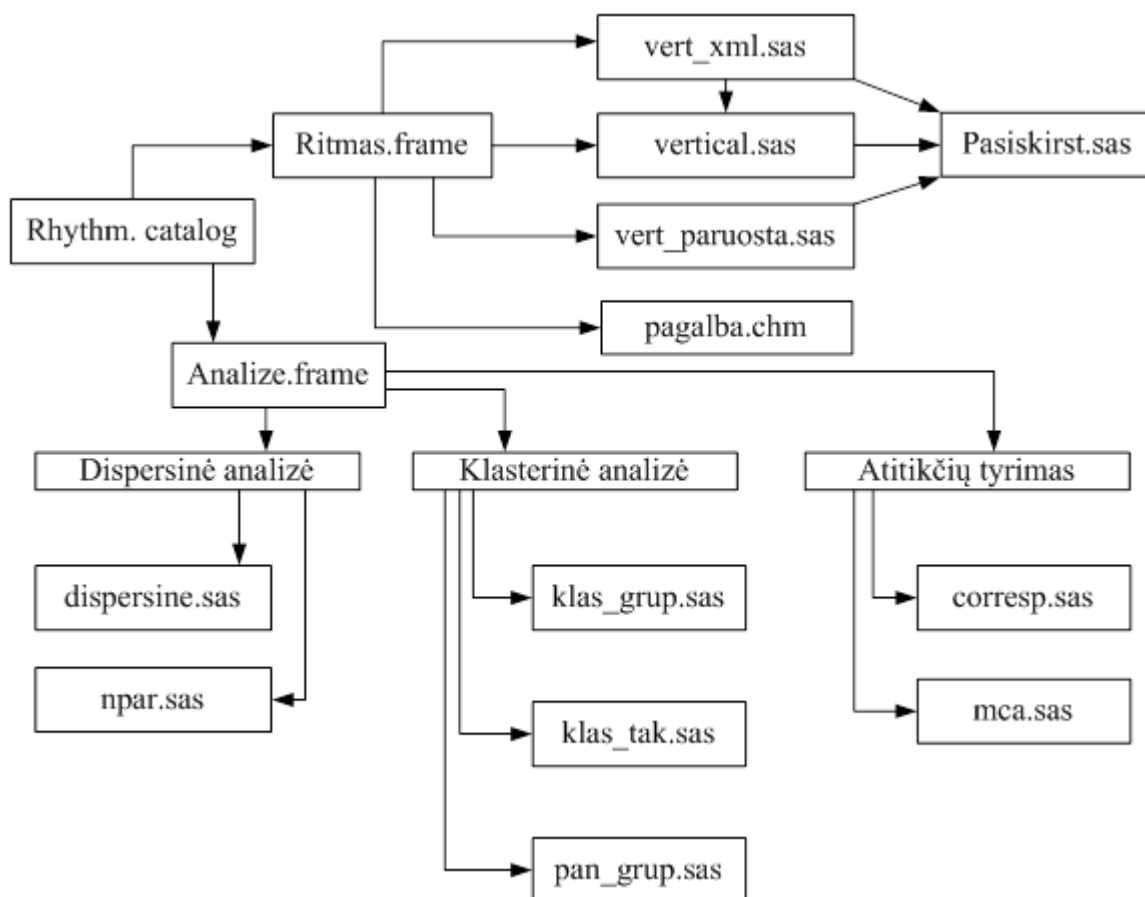
### 3 PROGRAMINĖ REALIZACIJA IR INSTRUKCIJA VARTOTOJUI

Programos vartotojo sąsają sudaro SAS katalogas, kuriame yra dvi programos langų formos ir du kiekvieną formą atitinkantys SCL programos kodai (žr. Pav. 3.1).



**Pav. 3.1** Programos katalogo struktūra

Programos atidarymo metu įvykdomas programos kodas (žr. priedą nr.1), kuriuo sukuriamos pagrindinės bibliotekos, kuriose saugomi darbiniai *sas* failai ir vartotojo sąsajos katalogas. Sukuriami makro kintamieji, naudojami kaip sutrumpinimai dokumentų atidarymo ir saugojimo procedūrose, įvykdomas vartotojo sąsajos katalogo paketo kompiliavimas.



**Pav. 3.2** Metodų įgyvendinimo struktūra programoje

Programą sudaro du langai – duomenų apdorojimo ir lyginamosios analizės. Duomenų apdorojimas ir modelių įgyvendinimas įvykdomas kreipiantis į kiekvienam metodui sukurtą *sas* failą, kreipiniai į failus yra programos formų SCL failuose. Programoje įgyvendinama dispersinė, klasterinė ir atitikties analizė (žr. Pav. 3.2). Kiekvienam analizės modeliui sukurtas atskiras *analize.frame* objektas „Kortelė“ (angl. *tab*). „Kortelėse“ arba skiltyse yra trumpi komentarai komponentėse *label*, galimi metodų pasirinkimai komponentėse *radiobox*. Sukurtas vartotojo vadovas saugomas faile „pagalba.pdf“, prieinamas pirmajame programos lange.

### 3.1 DUOMENŲ ĮKĖLIMAS IR APDOROJIMAS

Duomenų apdorojimas SAS aplinkoje sukurtoje vartotojo sąsajoje įgyvendinamas trims skirtingiems pradinių duomenų lygiams:

- neapdirbti duomenys, kurie įkeliami iš *xml* tipo failo;
- ritmo chromosomų sąrašas, sugrupuotas pagal kūrinių grupę ar pavadinimą, požymis – „grupe“;
- paruošti vertikalaus pjūvio dažniai, sugrupuoti pagal požymį „grupe“.

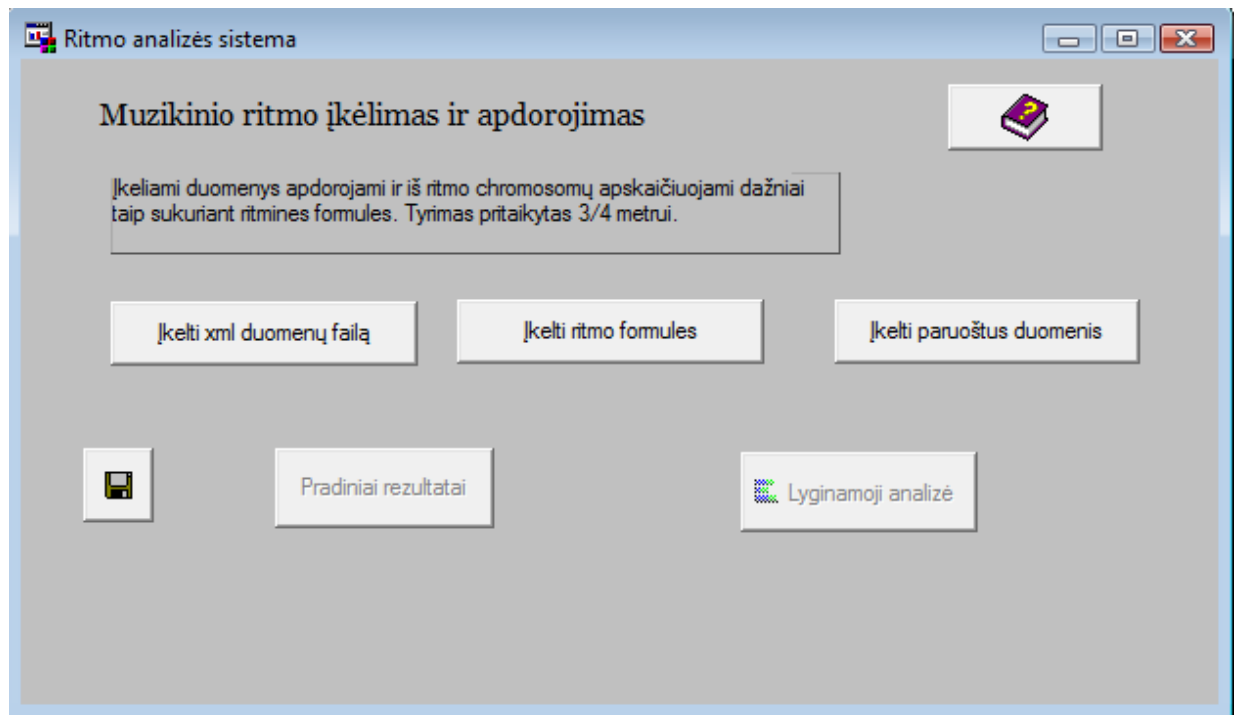
Kadangi 2 ir 3 atvejais duomenys įkeliami iš *xls* tipo failo, pirmojo lapo pavadinimas turėtų būti „Sheet1“, o duomenų failas, pavyzdžiui 3 atveju, turėtų atrodyti kaip pateikta lentelėje Lentelė 3.1. Kadangi programa įgyvendinta 3/4 metre ir visi rezultatų pavyzdžiai bus pateikti šiame metre. Tyrime 3/4 ir 6/8 metrai laikomi lygiaverčiais, nes natų kiekis jų taktuose yra vienodas.

**Lentelė 3.1 Apdorotų vertikalaus pjūvio duomenų pavyzdys**

GRUPE	_01	_02	_03	_04	...	_12
ciurlionio_ankstvy	90,58442		32,79221	6,493506	...	12,98701
ciurlionio_velyv	83,02387	12,99735	41,64456	14,58886	...	12,20159
...	...	...	...	...	...	...

Įkeliant duomenis iš *xml* failo neatmetama galimybė, kad *xml* žemėlapis neatitiks turimo failo, nors muzikinių *xml* failų argumentai yra vienodi, priklausomai nuo failo kokybės gali būti skirtinga reikšmių aprašymo struktūra. Šiuo atveju failai sukurti muzikinės programos *finale* pagalba, kuri naudojama muzikinių failų kūrimui ir redagavimui.

Saugojimo mygtuku galimas duomenų eksportavimas į *xls* tipo failą, eksportuojamas galutinis *sas* failas „vertikalė“ esantis *v* bibliotekoje programos aplanke *sas\_failai* (žr. Pav. 3.3).



**Pav. 3.3 Duomenų įkėlimo ir apdorojimo programos langas**

Įkėlus duomenis iš *xml* tipo failo atliekamas duomenų apdorojimas ir surašytos chromosomos, jeigu *xml* faile nėra klaidų, eksportuojamos į failą „papildomas.xls“ aplanke *Duomenys*. Jeigu yra klaidų taktuose, jų numeriai surašyti *sas\_failai* aplanke byloje *v.klaidos*. „Papildomas.xls“ byloje atlikus klaidingų taktų įvedimą rankiniu būdu, šis failas gali būti įkeliamas tolimesniam duomenų apdorojimui mygtuku *Įkelti ritmo formules*. Šiame žingsnyje duomenys apdorojami analizei – apskaičiuojami dažniai vertikaliai duomenų pjūviui.

Duomenų apdorojimas įgyvendinamas SAS<sup>®</sup> procedūrų pagalba:

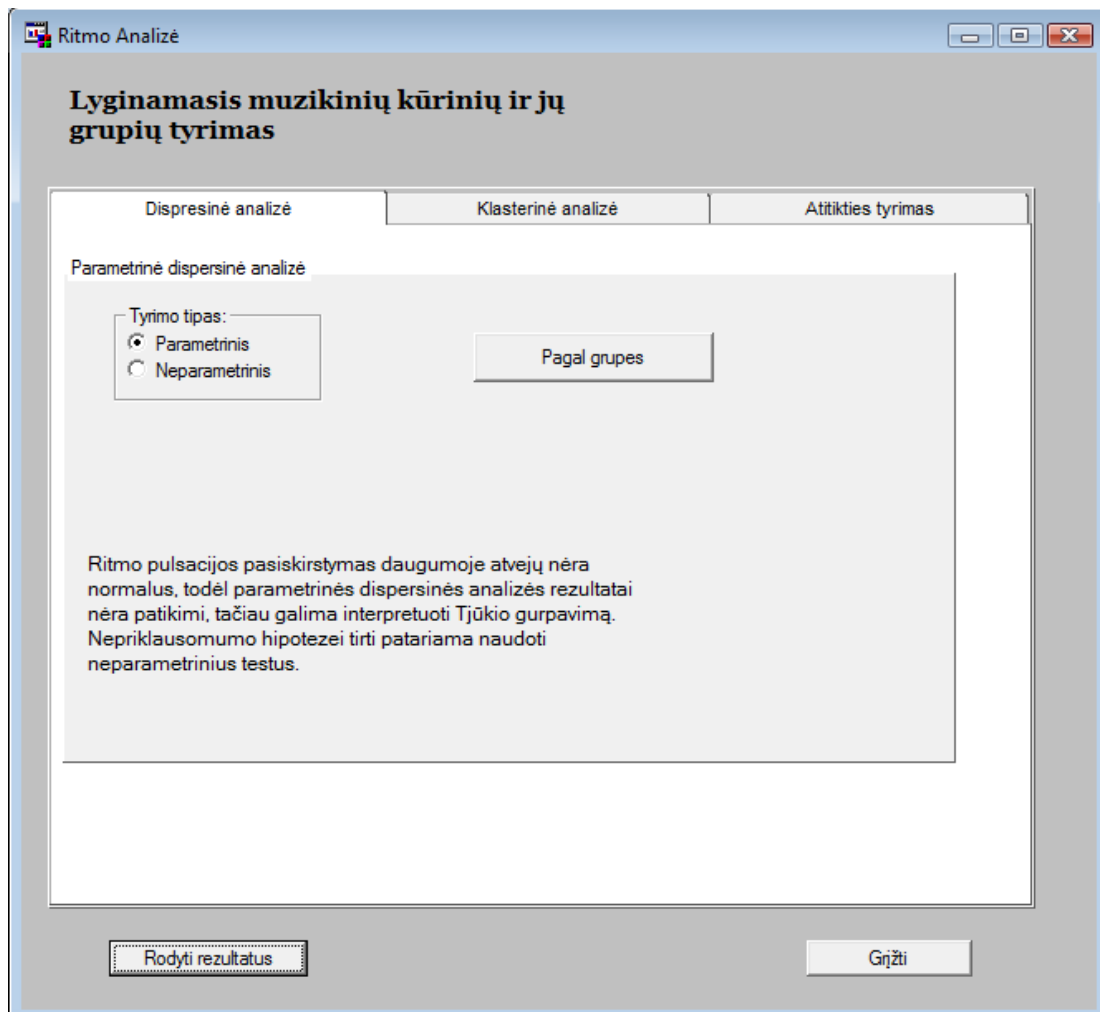
- *Import* – duomenų įkėlimui iš *xls* failų,
- *Export* – duomenų išsaugojimui *xls* failuose,
- *Transpose* – SAS failo duomenų matricos transponavimui,
- *Sort* – duomenų rikiavimui,
- *Freq* – dažnių skaičiavimui pagal grupes,
- *Summary* – dažnių sumavimui duomenų vertikaliai pjūviui.

Duomenų apdorojimo programiniai kodai pateikti priede Nr.4. Gautiems dažnių vektoriams mygtuko *Pradiniai rezultatai* paspaudimu pateikiamos aprašomosios charakteristikos, taktų kirčių histograma ir sugrupuota pasiskirstymo diagrama, pagal grupes. Aprašomosios charakteristikos su histograma pateikiamos SAS Base procedūros *univariate* pagalba. Taktų kirčių pasiskirstymo diagrama sugrupuota pagal kūrinių grupes pateikiama SAS/GRAPH *gchart* procedūros pagalba. Svarbu tai, kad mygtuko *Pradiniai rezultatai*

paspaudimu sukuriamas transponuotų duomenų failas, kuris vėliau naudojamas lyginamojoje duomenų analizėje, bei programos aplanke *Rezultatai* išsaugomas gautų rezultatų failas „pasiskirstymas.html“.

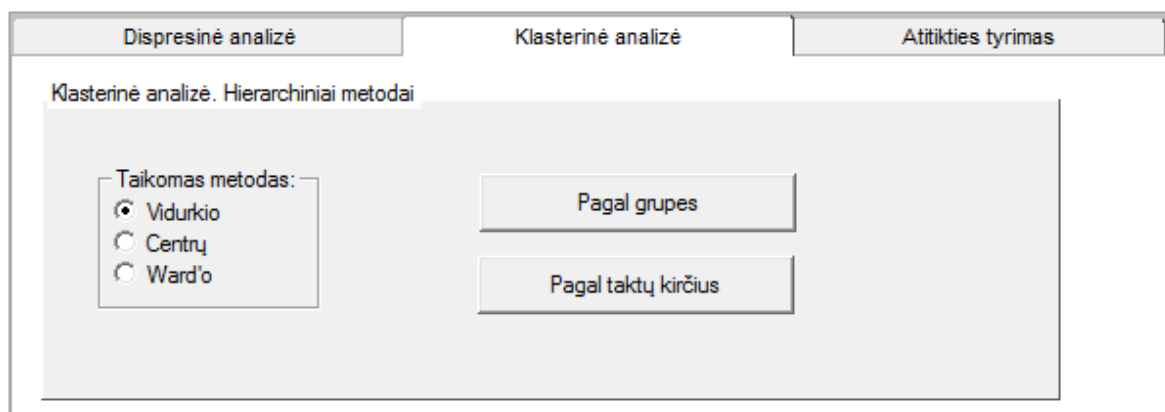
## 3.2 LYGINAMOJI RITMO ANALIZĖ

Mygtuko *Lyginamoji analizė* paspaudimu atidaromas antrasis programos langas (žr. Pav. 3.4). Vertikalaus pjūvio duomenų nepriklausomumui tirti pagal duomenų grupes arba takto kirčius yra du būdai – parametrinė dispersinė analizė arba neparametriniai kriterijai. Šiuo atveju svarbu tai, kad įkeliamų duomenų dažniai, kitaip nei kituose metoduose, turi būti nenormuoti, t.y. svarbu, kad reikšmės būtų didesnės už 1 ir nebūtų įvertintos procentais ar proporcija, *freq* procedūroje šie dažniai būtų vadinami *frequency count*. Parametrinė dispersinė analizė pateikia Fišerio statistiką ir jos tarpinius skaičiavimus, hipotezės priėmimo ir patikimumo parametrus ir Tjukio grupavimo rezultatus. Taip pat pateikiami grafikai skirti prielaidų patikrinimui. Rezultatai automatiškai išsaugomi faile „dispersine\_grup.html“. Pasirinkti neparametriniai kriterijai, atliekami nepriklausomumo hipotezės tikrinimui procedūroje *npar1way* yra *wilcoxon*, *median*, *savage* ir *vw* (Van der Waerdeno koef.). Šiuose rezultatuose, išsaugomuose faile „npar.html“, pateikiamos tarpinių koeficientų kiekvienai kūrinių grupei lentelės ir pagrindiniai rezultatai – chi-kvadrato statistika ir jos patikimumo tikimybė, leidžianti nuspręsti ar hipotezė priimta ar atmesta.



**Pav. 3.4 Lyginamosios Ritmo Analizės langas**

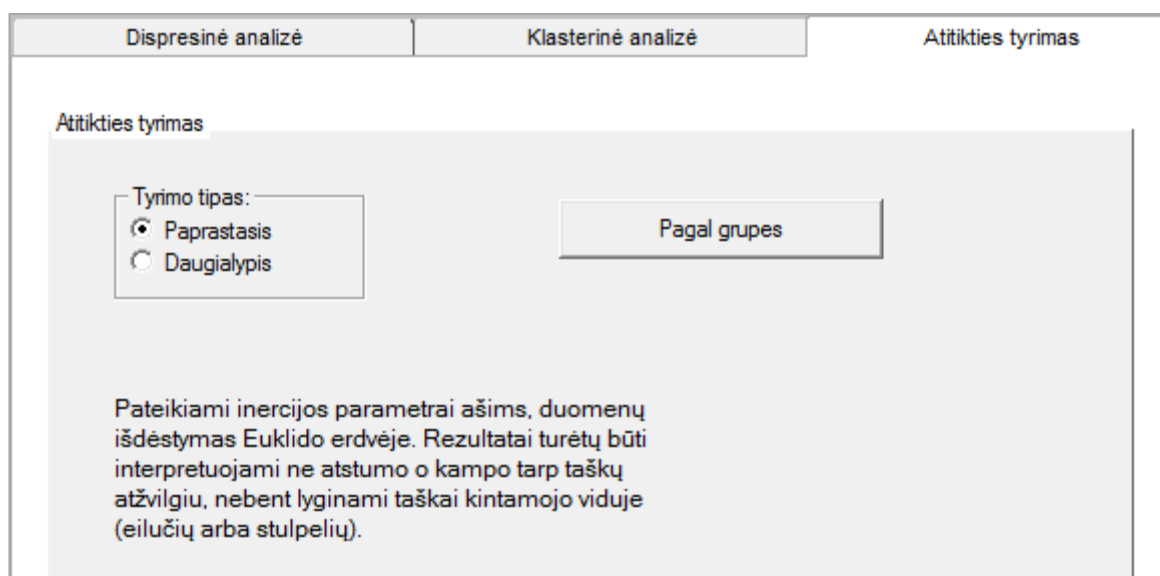
Antroje Analizės lango skiltyje įgyvendinamas atstumų ir panašumo koeficientų apskaičiavimas ir klasterinė analizė vienam atstumui ir vienam panašumo koeficientui, transformuotam į atstumą. Panašumo koeficientų apskaičiavimas įgyvendinamas tiriamoms duomenų grupėms, SAS/STAT procedūroje *distance* taikant metodus *Euclid*, *shape*, *simratio*, *corr*. Gautos trikampės koeficientų matricos saugomos faile „panašumo\_koef.html“. Pagal *Euclid* ir *dcorr* metodus gautus atstumus įgyvendinama hierarchinė klasterinė analizė naudojant SAS/STAT procedūras *cluster* ir *tree*, grupavimo medžiui pavaizduoti. Klasterinė analizė gali būti įgyvendinama pagal duomenų grupes arba pagal takto kirčius atitinkamo mygtuko paspaudimu (žr. Pav. 3.5), taip pat yra galimybė pasirinkti klasterių apjungimo metodą: vidurkio, centrų arba Ward'o (atitinkamai SAS procedūroje *Average*, *Centroid*, *Ward*). Rezultatuose pateikiamos ir *pseudo* statistikos, kurios gali padėti interpretuoti dendrogramą. Grupių klasterizavimo rezultatai saugomi faile „klasterine\_grup.html“, jų grupavimo paveikslai saugomi failuose „klaster1.gif“, „klaster2.gif“, taktu kirčių klasterizavimo rezultatai saugomi – „klasterine\_taktu.html“, „klaster3.gif“, „klaster4.gif“ failuose.



**Pav. 3.5 Klasterinės analizės skiltis Analizės lange**

Papildomai sukurtas klasterinės analizės failas „klast\_grup\_bandymams.sas“ saugomas aplanke *Kodai*, jei vartotojas norėtų išbandyti kitus SAS procedūroje *distance* galimus atstumus ir juos palyginti. Kulback-Leibler ir simetrinė Jeffreys divergencijų skaičiavimai nėra įgyvendinti šioje programinėje įrangoje.

Atitikties tyrimas programoje įgyvendinamas SAS/STAT procedūra *corresp*, atitikties žemėlapiui pateikiami makro komandos *%plotit* pagalba. Galimi du tyrimo būdai: paprastasis ir daugialypis (žr. Pav. 3.6). Paprastojo atitikties tyrimo atveju pateikiamos visos rezultatų lentelės: inercijos ir chi-kvadrato reikšmių, eilučių ir stulpelių koordinatų, eilučių ir stulpelių taškus aprašančios statistikos. Paprastasis atitikties žemėlapis, interpretuojamas inercijos, kampo tarp taškų, atstumo tarp taškų kintamojo viduje atžvilgiu.



**Pav. 3.6 Atitikties tyrimo skiltis Analizės lange**

Daugialypio atitikties tyrimo atveju pateikiamos dimensijų inercijos ir chi-kvadrato reikšmių, stulpelių koordinatų pirmose dviejose dimensijose lentelės, taškų įvertinimas

proporcijomis inercijos atžvilgiu ir kokybės parametru dimensijų atžvilgiu. Rezultatai saugomi faile „atitikciu\_analize.html“ ir „atitikciu.gif“ žemėlapiu paveikslėlis.

Daugialypis atitikties tyrimas įgyvendinamas *corresp* procedūroje nurodant pasirinktis *mca*, pateikiamos analogiškos ataskaitos kaip ir paprastuoju atveju. Daugialypio atitikties tyrimo rezultatai gali būti interpretuojami taip pat kaip ir paprastojo, tačiau dažniausiai jų inercija būna maža. MCA rezultatai ir žemėlapis saugomi failuose „mca\_rezultatai.html“ ir „mca.gif“.

Norėdamas keisti metodų rezultatų išvestis, vartotojas turėtų papildyti arba taisyti naudojamus *sas* failus patalpintus aplanke *Kodai*. Norėdamas atlikti tyrimą su kitais duomenimis, vartotojas gali grįžti į pradinį langą (mygtukas *Grįžti*) ir įkelti naują duomenų failą tyrimui. Mygtuko *Rodyti rezultatus* paspaudimu atidaroma papildoma sekcija, kurioje matoma visų įvykdytų tyrimų rezultatų istorija. Pagalbos mygtuku atidaromas programos naudotojo vadovas.



## 4 DISKUSIJOS

### 4.1 PASISKIRSTYMO PALYGINIMAS IR NEPRIKLAUSOMUMO TYRIMAS

Čiurlionio fortepijoninės muzikos kūriniai pasirinkti atsitiktinai, siekiant kad būtų analizuojama daugiau muzikinių žanrų. Pagal sukūrimo metus, jie priklauso Čiurlionio vėlyvajai kūrybai, kaip ir pasirinktas papildomas kūrinys – Preljudas B-dur iš simb. eilėgarsių II.

Pastebėta, kad interpretuoti taktų pulsacijos histogramas tarpusavyje yra patogiau nei kiekvieną takto kirtį atskirai visose duomenų grupėse. Lyginant kūrinių grupių vektorius dažnių grafike, labiausiai netolygia pulsacija pasižymi simfoninės poemos „Jūra“ ritmas, net silpnosios takto dalys pasiskirsčiusios netolygiai. Sudarykime tris taktų kirčių lygius pagal stiprumą – nuo stipriausio iki silpniausio:

1. 1, 5, 9;
2. 3, 7, 11;
3. 2, 4, 6, 8, 10, 12.

Lyginant pasiskirstymą kiekvieno lygio atžvilgiu atskirai, pastebime, kad fortepijoninės muzikos taktų kirčių histogramos kiekvienas stiprumo lygis yra pasiskirstęs tolygiai. Dzūkų liaudies muzikos trečiasis takto kirčių lygis yra pasiskirstęs tolygiai. Jei lyginsime pirmojo ir antrojo lygio atžvilgiu, dzūkų liaudies ir Čiurlionio vėlyvosios kūrybos histogramos yra panašios formos atžvilgiu. Silpnosios takto dalys, Čiurlionio vėlyvojoje, ankstyvojoje ir fortepijoninėje kūryboje šiuo atveju pasižymi didesniais dažniais nei lenkų ir dzūkų liaudies muzikoje. D.Kučinsko ir R.Ambrazevičiaus tyrimuose lyginta takto dalių pulsacija: pastebėtas silpnosios takto dalies stiprumas, kuris yra išskirtinis lyginant Čiurlionio ankstyvąją ir vėlyvąją kūrybą su lenkų ir dzūkų liaudies muzika. Pateiktame tyrime pastebėta, kad ir simfoninės poemos „Jūra“ 3/4 metro ritmas pasižymi šia savybe.

Nepriklausomumo tyrimui taikomos dispersinės analizė rezultatai nėra patikimi, nes netenkinama prielaidos apie analizuojamo dydžio normalųjį pasiskirstymą, taip pat pažeidžiamos ir kitos prielaidos apie paklaidų pasiskirstymą, remiantis apibrėžtumo koeficientu šio modelio rezultatai paaiškina tik 1.46% duomenų sklaidos. Tjukio grupavimą galime interpretuoti tačiau rezultatai bus nepatikimi, nes atmetama prielaida apie dispersijų homogeniškumą. Šiuo atveju kūriniai grupuojami į Čiurlionio ir liaudies kūrybos grupes. Taip pat kitu atveju su Čiurlionio kūryba grupuojama dzūkų liaudies muzika.

Neparametrinių kriterijų atveju gauti analogiški rezultatai – taktų kirčių pulsacija yra priklausoma nuo kūrinių grupių.

## 4.2 PANAŠUMO KOEFICIENTAI IR KLASTERINĖ ANALIZĖ

J-divergencijos rezultatai stipresniųjų takto dalių atžvilgiu (jei lygintume aštuntinės natos tikslumu) analogiškai Tjukio grupavimo rezultatams (nors jie yra nepatikimi) – simfoninė poema „Jūra“ grupuojama su lenkų liaudies muzika, tiek vienu iš atvejų pradinių duomenų grupių palyginime, tiek papildžius duomenis Preljudu B-dur. Šio papildomo kūrinio J-divergencija su tiriamomis grupėmis išskiria stipriausią panašumą vėlyvajai Čiurlionio kūrybai, kuriai jis ir priklauso, tačiau artumu fortepijoninei muzikai nepasižymi lyginant su kitų grupių atstumais. Jeigu išrikiuotume gautas grupes panašumo koeficiento mažėjimo tvarka, antra artimiausia Preljudui kūrinių grupė būtų dzūkų liaudies muzikos. Kitu atveju, kai skaičiuojama divergencija fortepijoninei muzikai, gaunamos dvi artimiausios grupės iš tiriamų – Čiurlionio vėlyvoji muzika su atstumu 0.089 ir papildomas kūrinys Preljudas su atstumu 0.127. Matome, kad gauti rezultatai teisingi stipriųjų takto dalių atžvilgiu.

Kitų atstumų skaičiavimais gaunami analogiškai rezultatai išskyrus koreliacijos koeficiento atveju. Euklido, formos ir panašumo atstumų rezultatai klasterių dendrogramoje išskiria dvi pradines grupes – liaudies muzikos ir Čiurlionio vėlyvosios, fortepijoninės muzikos. Koreliacijos koeficiento paversto atstumu atveju grupavimas vyksta kas žingsnį prijungiant po vieną kūrinių grupę prie pradinės, kurią sudaro dzūkų liaudies ir Čiurlionio vėlyvoji muzika. Tačiau šiuo atveju galėtume sakyti kad grupavimas nėra teisingas turimų duomenų atžvilgiu, nes fortepijoninė muzika nėra gretinama su vėlyvąja kūryba. Šis atstumas duoda gerus rezultatus muzikiniu aspektu taktų kirčių grupavime, kur sugrupuojamos silpnos ir stiprios pulsacijos takto dalys.

## 4.3 ATITIKTIES TYRIMAS

Įvertinus koordinačių kokybę ir inercijos dalį kurią sudaro kiekviena duomenų grupė, matome, kad ankstyvosios Čiurlionio kūrybos taškas nėra patikimas. Simfoninės poemos taškas sudaro didelę dalį inercijos kiekvienoje dimensijoje, tačiau jis neturi artimų taškų. Pirmoji dimensija sudaro 55.34% visos inercijos. Remdamiesi taškų inercija ašiai ir jų koreliacija su ašimi, galime teigti, kad pirmoji dimensija yra faktorius apibūdinantis Čiurlionio vėlyvąją kūrybą - artimiausi taškai yra vėlyvosios ir fortepijoninės muzikos grupės. Antroji ašis apibūdina lenkų ir dzūkų liaudies muziką, kurių taškai taip pat labiausiai koreliuoti su ašimi bei pasižymi mažu atstumu tarpusavy. Tačiau ši ašis apibūdina mažesnę dalį inercijos – 29.53%.

Papildomas kūrinys Preliudas šioje analizėje yra artimiausias liaudies muzikai ir kartu priklauso antrajam faktoriui, jei vertinsime atstumą – artimesnis dzūkų liaudies muzikai.

Pastebėta, kad šių duomenų tyrimui daugialypė atitikties analizė nėra reikalinga, nes paprastosios analizės atveju rezultatai pakankamai kokybiški ir nesunkiai interpretuojami. Ją patartina naudoti, kai norima palyginti taškus tarp stulpelių ir eilučių kintamųjų.

#### **4.4 REZULTATŲ PALYGINIMAS**

Naudotų metodų duomenų grupių artumui įvertinti rezultatai panašūs, tačiau reikšmės nėra apribotos uždaru intervalu, todėl vertinama turimų duomenų grupėje, t.y. nežinome ar galėtume šiuos rezultatus taikyti populiacijos atžvilgiu, nes negalime įvertinti ryšio tarp tiriamų grupių stiprumo. Dėl šios priežasties negalime nustatyti panašumo naudodami vieną metodą. Apibendrinus šiuos rezultatus, matome kad nei vienas iš pateiktų metodų nedavė klaidingų rezultatų, fortepijoninės muzikos kūriniai ir papildomas Preliudas daugumoje rezultatų buvo priskirti tai kryptčiai kuriai priklauso, bei kai kuriais atvejais buvo gretinami su dzūkų liaudies muzika. Simfoninė poema „Jūra“ pasižymėjo išskirtinumu beveik visuose metodų rezultatuose. Ji dažniausiai buvo gretinama su lenkų liaudies muzika.

Horizontalaus duomenų pjūvio atveju  $3/4$  metre, buvo atmesta dalis labai retai pasikartojančių ritmų. Šio pjūvio lyginamojoje analizėje į klasterius sugrupuota: Čiurlionio vėlyvoji su dzūkų liaudies muzika, lenkų liaudies muzika su simfonine poema „Jūra“. Analogiški rezultatai gauti taikant skirtingus panašumo koeficientus ir atstumus.

Tirtos duomenų grupės buvo artimos laikmečiu ir muzikiniu aspektu, dėl to jų palyginimas yra tikslingas.

## IŠVADOS

- Tyrimui taikytas vertikalus duomenų pjūvis yra vaizdesnis ir patikimesnis taikant statistinius metodus.
- Programinė įranga sukurta su vartotojo sąsaja skirta kokybiškų muzikinių *xml* failų įkėlimui, apdorojimui ir lyginamajai analizei. Analizės modeliai gali būti nesudėtingai modifikuojami, nes kiekvienas jų saugojamas atskirame *sas* faile.
- Analizuojant tam tikrų ritmo elementų atžvilgiu arba tikslumu, gauti reikšmingi rezultatai, neprarastos duomenų kategorijų savybės.
  - Ritmo pulsacijos arba vektorių dvejetainis interpretavimas leido taikyti daugiau metodų.
  - Panašumo koeficientų, atstumų, koeficientų paverstų atstumais, divergencijų rezultatai išskyrė pagrindines duomenų grupes, bei klasterizavimo pagalba buvo išryškinti koeficientų skirtumai ir jų taikymo galimybės.
  - Atitikties tyrimo rezultatai sutapo su klasterinės analizės rezultatais – išskirtos tokios pačios duomenų grupės.
  - Atitikties tyrime išskirti du pagrindiniai faktoriai, kiekvienas faktorius apibūdina porą duomenų grupių. Šie rezultatai informatyvūs grupavimo ir kokybiški analizės parametrų atžvilgiu.
- Naudotam duomenų pjūviui bandyta taikyti transformacijas ir pakeitimus, su tikslu taikyti parametrinius metodus. Tačiau negavus reikiamo rezultato, buvo taikomi neparametriniai metodai, skirti kategoriniams duomenims, pavyzdžiui faktorinės analizės atitikmuo – atitikties analizė.
- Taikant neparametrinius metodus ar atstumus, rezultatams įvertinti įtraukta daugiau panašių grupių duomenų. Palyginamosios analizės rezultatai dėl to nepakito, tik kai kuriais atvejais sumažėjo jų reikšmingumas.
- Lyginamosios ritmo analizės rezultatai rodo, kad esant teisingam duomenų grupavimui tai patvirtina bent pora naudotų metodų.

## REKOMENDACIJOS

- Ritmo formulių užrašymas binariniu kodu sutraukia skirtingus ritmus į tam tikras grupes. Norint analizuoti ritmą atsižvelgiant į tai iš kokio ilgio natų jis susideda, galima sukurti skaitinį kodą, kuriame atspindėtų natų ir pauzių ilgiai.
- Atliktame tyrime didžiąją dalį užima neparametriniai metodai. Tyrimo eigoje bandyta pritaikyti įvairias duomenų transformacijas, kad būtų galima taikyti parametrinius metodus, arba parametrų ribų įvertinimą, reikalaujantį normaliojo pasiskirstymo. Tyrime tai įgyvendinti nepavyko dėl pasirinkto analizuoti duomenų pjūvio, tačiau rekomenduotina rasti kitokių duomenų skaidymo ar transformavimo būdų, nes parametriniai metodai yra informatyvesni.
- Šiame tyrime nėra išnaudotos visos atitikties tyrimo galimybės – kampo tarp kintamųjų atžvilgiu yra galimybė įvertinti kaip koreliuoti taškai tarp skirtingų kintamųjų. Tai būtų galima padaryti, pavyzdžiui, tiriant duomenų grupių koreliaciją su silpnųjų ir stipriųjų taktų kirčių grupėmis.
- Tyrimų eigoje buvo išbandyti genetiniai algoritmai, ritmo chromosomų atkūrimui. Tyrimas nebuvo išplėtotas, nes juo sprendžiama generavimo problema, tačiau suradus tinkamą tikslo funkciją ar mutacijos parametrus būtų galima pritaikyti šiuos algoritmus trūkstamų kūrinų dalių atkūrimui.
- Vienas siekinių būtų rasti arba pritaikyti parametrinius metodus, kurie įvertintų panašumo lygį arba stiprumą skaitine reikšme, kaip pavyzdžiui koreliacijos koeficiento arba KL-divergencijos Markovo grandinėms atveju.

## **PADĖKOS**

Dėkoju:

- Lietuvos Mokslo Tarybai už suteiktą galimybę dalyvauti studentų mokslinėje praktikoje;
- Philip Morris Baltic už paramą mokslinės praktikos tęsimui;
- Prof.dr. Dariui Kučinskui už suteiktas galimybes ir pagalbą tyrimų įgyvendinime;
- Dr. Tomui Ruzgui už patarimus ir pagalbą įgyvendinant statistinius tyrimus.

## LITERATŪRA

1. *Analysis of Variance*. <http://www.sjsu.edu/faculty/gerstman/StatPrimer/anova-a.pdf> (žiūrėta: 2011-01-15).
2. Balčikonis, J. *Druskininkų dainos*. Vilnius: Vaga, 1972.
3. Benko, Michal ir Lejeune, Michel. "Correspondence analysis." p.21.
4. Cha, Sung-Hyuk. "Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions." *INTERNATIONAL JOURNAL OF MATHEMATICAL MODELS AND METHODS IN APPLIED SCIENCES* vol.1, no. 4 (2007): p.300-307.
5. Creators of Statistica, data analysis software. *Correspondence analysis*. <http://www.statsoft.com/textbook/correspondence-analysis/> (žiūrėta: 2011-01-12).
6. Četkauskaitė, G. *Dzūkų melodijos*. Vilnius: Vaga, 1981.
7. Čiurlionis, Mikalojus Konstantinas. *Kūriniai fortepijonui*. Kaunas: J.Petronis, 2004.
8. Dahlig-Turek, Ewa. "Studying Rhythm Morphology." *Lietuvos muzikologija*, 2009: p.127-137.
9. Friendly, Michael. "A Reader's Guide to Visualizing Categorical Data." *26 Annual SAS® Users Group International Conference*. Cary, NC: SAS Institute Inc., 2001. p.173-26.
10. Garson, David. *Correspondence analysis*. 2008. <http://faculty.chass.ncsu.edu/garson/PA765/correspondence.htm>. (žiūrėta: 2011-02-01).
11. Greenacre, Michael. "Appendix A (Theory of Correspondence Analysis)." <http://statmath.wu.ac.at/courses/CAandRelMeth/> (žiūrėta: 2011-01-15).
12. Yeo, David. *Applied Clustering Techniques Course Notes*. JAV: SAS Institute Inc., 2003.
13. Kašponis, Rimtautas. *Lietuvių muzikos melodika ir harmonija*. 1992.
14. Kumar, Pranesh. "Generalized Relative J-Divergence Measure." *Int. J. Contemp. Math. Sci.* Vol.1, no.13 (2006): p.597 - 609.
15. Liu, Yi-Wen. "Modeling Music as Markov Chains:." Music 254 Final report, 2006.
16. Lu, David. "Automatic Music Transcription Using Genetic." 2006.
17. Micheloud, François-Xavier. *Correspondence Analysis*. 1997.
18. Miranda, Eduardo Reck. *Composing Music with Computers*. Vol. ISBN 0 240 51567 6. Burlington: Focal Press, Elsevier, 2002.
19. Narayanan, Anupama and Watts, Donna. *Exact Methods in the npar1way procedure*. Cary, NC: SAS Institute Inc.
20. Pappas, Paul A. ir DePuy, Venita. *An Overview of Non-parametric Tests in SAS: When, Why and How*. Paper TU04, Durham, JAV: Duke Clinical Research Institute, 2004.

21. SAS Institute, Inc. *SAS/STAT® User's Guide*. Vol. Version 8. Cary, NC: SAS Institute Inc., 1999.
22. Schoner, Jeffrey. "Iannis Xenakis' Musique Stochastique: System Design and." *Seminar Computergestützte Komposition, WS*. Universität des Saarlandes, 2003.
23. *Scores for Linear Rank and One-Way ANOVA Tests*. [http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug\\_npar1way\\_a0000000200.htm](http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_npar1way_a0000000200.htm) (žiūrėta: 2011-01-10).
24. Shlens, Jonathon. *Notes on Kullback-Leibler Divergence and Likelihood Theory*. Systems Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 9203, 2007.
25. Unesco, Idams statistical software. *Correspondence analysis*. [http://www.unesco.org/webworld/idams/advguide/Chapt6\\_5.htm](http://www.unesco.org/webworld/idams/advguide/Chapt6_5.htm) (žiūrėta: 2011-01-20).



## PRIEDAS NR.1 PRADINIS PROGRAMOS KODAS IR SCL FAILAI

### Failas prad.sas.

```
libname muz "%sysget(pagrindinis)";
%let place=C:\Users\HT\Desktop\RITMO ANALIZE\XML failai;
%let exelio=C:\Users\HT\Desktop\RITMO ANALIZE\Duomenys;
%let rez=C:\Users\HT\Desktop\RITMO ANALIZE\Rezultatai;
libname v "C:\Users\HT\Desktop\RITMO ANALIZE\sas_failai";

proc build catalog=muz.rhythm batch;
compile select=(Ritmas.frame Analyze.frame);
run;
```

### Failas Ritmas.scl

```
INIT:
control always;
RETURN;
PUSHBUTTON5:
call display('Analyze.frame');
RETURN;
PUSHBUTTON1:
rc=FILEDIALOG('OPEN',failas,'','&place','*.xml');
IF failas^="" THEN DO;
    call symput('path', failas);
    call execcmd('submit file = "C:\Users\HT\Desktop\RITMO ANALIZE\Kodai\vert_xml.sas"');
pushbutton3.enabled="Yes";
pushbutton5.enabled="Yes";
END;
RETURN;
PUSHBUTTON2:
rg=FILEDIALOG('OPEN',failas1,'','&exelio','*.xls');
IF failas1^="" THEN DO;
    call symput('path', failas1);
    call execcmd('submit file = "C:\Users\HT\Desktop\RITMO ANALIZE\Kodai\vertikal.sas"');
pushbutton3.enabled="Yes";
pushbutton5.enabled="Yes";
END;
RETURN;
PUSHBUTTON8:
rc=FILEDIALOG('OPEN',failas,'','&exelio','*.xls');
IF failas^="" THEN DO;
    call symput('path', failas);
    call execcmd('submit file = "C:\Users\HT\Desktop\RITMO ANALIZE\Kodai\vert_paruosta.sas"');
END;
pushbutton3.enabled="Yes";
```

```

pushbutton5.enabled="Yes";

RETURN;

PUSHBUTTON6:

rs=FILEDIALOG('saveas',saugot,'','C:\Users\WHT\Desktop\WRITMO ANALIZE\Rezultatai.xls','*.xls');
IF saugot^="" THEN DO;
    call symput('path', saugot);
    call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\exp.sas"');
END;

RETURN;

PUSHBUTTON3:

call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\pasiskirst.sas"');

RETURN;

PUSHBUTTON4:

call execcmd('wbrowse "C:\Users\WHT\Desktop\WRITMO ANALIZE\pagalba.pdf" ');

RETURN;

```

### Failas Analyze.scl

```

INIT:

control always;
Radiobox1.selectedItem="Parametrinis";
Radiobox5.selectedItem="Vidurkio";
Radiobox2.selectedItem="Paprastasis";

RETURN;

MAIN:

if Radiobox5.selectedItem = "Vidurkio" then call symput('met','average');
if Radiobox5.selectedItem = "Ward'o" then call symput('met','ward');
if Radiobox5.selectedItem = "Centru" then call symput('met','centroid');

RETURN;

PUSHBUTTON2:

if Radiobox1.selectedItem="Parametrinis" then do;
    call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\dispersine.sas"');
end;
else do;
    call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\wpar.sas"');
end;

RETURN;

PUSHBUTTON3:

call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\klast_grup.sas"');

RETURN;

PUSHBUTTON1:

call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\klast_tak.sas"');

RETURN;

```

```
PUSHBUTTON4:
call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\pan_grup.sas"');
RETURN;

PUSHBUTTON5:
if Radiobox2.selectedItem="Paprastasis" then do;
    call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\corresp.sas"');
end;
else do;
    call execcmd('submit file = "C:\Users\WHT\Desktop\WRITMO ANALIZE\Kodai\mca.sas"');
end;
RETURN;
```

## PRIEDAS NR.2 XML ŽEMĒLAPIS

### Failas import.map

```
<?xml version="1.0" encoding="UTF-8"?>

<!-- ##### -->
<!-- 2003-04-08T15:03:16 -->
<!-- SAS XML Libname Engine Map -->
<!-- Generated by XML Mapper, 9.1.10.20030407.1378 -->
<!-- ##### -->

<SXLEMAP version="1.2" name="SXLEMAP">

  <!-- ##### -->

  <TABLE name="Natos">
    <TABLE-PATH syntax="XPath"/>/score-partwise/part/measure/note</TABLE-PATH>

    <COLUMN name="taktas" retain="YES" ordinal="YES">
      <INCREMENT-PATH syntax="XPath"/>/score-partwise/part/measure</INCREMENT-PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>integer</DATATYPE>
      <FORMAT width="4">Z</FORMAT>
    </COLUMN>

    <COLUMN name="metras" retain="YES">
      <PATH syntax="XPath"/>/score-partwise/part/measure/attributes/time/beats</PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>integer</DATATYPE>
      <FORMAT width="1">Z</FORMAT>
    </COLUMN>

    <COLUMN name="Type">
      <PATH syntax="XPath"/>/score-partwise/part/measure/note/type</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>11</LENGTH>
    </COLUMN>

    <COLUMN name="dot" retain="YES" ordinal="YES">
      <INCREMENT-PATH syntax="XPath"/>/score-partwise/part/measure/note/dot</INCREMENT-PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>integer</DATATYPE>
      <FORMAT width="3">Z</FORMAT>
    </COLUMN>

    <COLUMN name="tie" retain="YES" ordinal="YES">
      <INCREMENT-PATH          syntax="XPath"/>/score-partwise/part/measure/note/tie</INCREMENT-
PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>integer</DATATYPE>
      <FORMAT width="3">Z</FORMAT>
    </COLUMN>
```

```

<COLUMN name="step">
  <PATH syntax="XPath">/score-partwise/part/measure/note/pitch/step</PATH>
  <TYPE>character</TYPE>
  <DATATYPE>string</DATATYPE>
  <LENGTH>1</LENGTH>
</COLUMN>

<COLUMN name="octave">
  <PATH syntax="XPath">/score-partwise/part/measure/note/pitch/octave</PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>integer</DATATYPE>
  <LENGTH>1</LENGTH>
</COLUMN>

<COLUMN name="alter">
  <PATH syntax="XPath">/score-partwise/part/measure/note/pitch/alter</PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>integer</DATATYPE>
  <LENGTH>2</LENGTH>
</COLUMN>

<COLUMN name="chord" retain="YES" ordinal="YES">
  <INCREMENT-PATH syntax="XPath">/score-partwise/part/measure/note/chord</INCREMENT-
PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>integer</DATATYPE>
  <FORMAT width="3">Z</FORMAT>
</COLUMN>

<COLUMN name="grace" ordinal="YES">
  <INCREMENT-PATH syntax="XPath">/score-partwise/part/measure/note/grace</INCREMENT-
PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>integer</DATATYPE>
  <FORMAT width="3">Z</FORMAT>
</COLUMN>

<COLUMN name="actual-notes">
  <PATH syntax="XPath">/score-partwise/part/measure/note/time-modification/actual-
notes</PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>integer</DATATYPE>
</COLUMN>

<COLUMN name="normal-notes">
  <PATH syntax="XPath">/score-partwise/part/measure/note/time-modification/normal-
notes</PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>integer</DATATYPE>
</COLUMN>

```

```
<COLUMN name="slur" retain="YES" ordinal="YES">
  <INCREMENT-PATH                                syntax="XPath"/>/score-
partwise/part/measure/note/notations/slur</INCREMENT-PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>integer</DATATYPE>
  <FORMAT width="3">Z</FORMAT>
</COLUMN>

</TABLE>

</SXLEMAP>
```

## PRIEDAS NR.3 XML FAILO PAVYZDYS

### Muzikinio xml failo pirmojo takto aprašymo pavyzdys (iš 4 pjesės Rudenelis.xml)

```
<!------->
<part id="P1">
  <measure number="1" width="319">
    <print>
      <page-layout>
        <page-height>1642</page-height>
        <page-width>1161</page-width>
        <page-margins>
          <left-margin>83</left-margin>
          <right-margin>81</right-margin>
          <top-margin>72</top-margin>
          <bottom-margin>121</bottom-margin>
        </page-margins>
      </page-layout>
      <system-layout>
        <system-margins>
          <left-margin>60</left-margin>
          <right-margin>0</right-margin>
        </system-margins>
        <top-system-distance>211</top-system-distance>
      </system-layout>
    </print>
    <attributes>
      <divisions>8</divisions>
      <key>
        <fifths>-1</fifths>
        <mode>major</mode>
      </key>
      <time>
        <beats>3</beats>
        <beat-type>4</beat-type>
      </time>
      <staves>2</staves>
      <clef number="1">
        <sign>G</sign>
        <line>2</line>
      </clef>
      <clef number="2">
        <sign>F</sign>
        <line>4</line>
      </clef>
    </attributes>
    <sound tempo="120"/>
    <direction placement="below">
      <direction-type>
        <dynamics default-y="-83">
          <mf/>
        </dynamics>
      </direction-type>
    </direction>
  </measure>
</part>
```

```

        </direction-type>
        <staff>1</staff>
        <sound dynamics="83"/>
    </direction>
    <direction placement="above">
        <direction-type>
            <words default-y="35" font-size="13.7" font-weight="bold" relative-x="-
32">Agitato</words>
        </direction-type>
        <staff>1</staff>
    </direction>
    <note>
        <rest/>
        <duration>24</duration>
        <voice>1</voice>
        <staff>1</staff>
    </note>
    <direction placement="below">
        <direction-type>
            <words default-y="-84" font-size="12" font-style="italic" font-weight="bold">e
cresc.</words>
        </direction-type>
        <offset>-17</offset>
        <staff>1</staff>
    </direction>
    <backup>
        <duration>24</duration>
    </backup>
    <note default-x="108">
        <pitch>
            <step>A</step>
            <octave>2</octave>
        </pitch>
        <duration>4</duration>
        <voice>3</voice>
        <type size="cue">eighth</type>
        <stem default-y="-67">down</stem>
        <staff>2</staff>
        <beam number="1">begin</beam>
    </note>
    <note default-x="142">
        <pitch>
            <step>A</step>
            <octave>3</octave>
        </pitch>
        <duration>4</duration>
        <voice>3</voice>
        <type size="cue">eighth</type>
        <stem default-y="-67">down</stem>
        <staff>2</staff>
        <beam number="1">continue</beam>
    </note>
    <note default-x="176">

```



```

    <pitch>
      <step>A</step>
      <octave>2</octave>
    </pitch>
    <duration>4</duration>
    <voice>3</voice>
    <type size="cue">eighth</type>
    <stem default-y="-67">down</stem>
    <staff>2</staff>
    <beam number="1">continue</beam>
  </note>
  <note default-x="212">
    <pitch>
      <step>A</step>
      <alter>-1</alter>
      <octave>3</octave>
    </pitch>
    <duration>4</duration>
    <voice>3</voice>
    <type size="cue">eighth</type>
    <accidental>flat</accidental>
    <stem default-y="-67">down</stem>
    <staff>2</staff>
    <beam number="1">continue</beam>
  </note>
  <note default-x="250">
    <pitch>
      <step>A</step>
      <octave>2</octave>
    </pitch>
    <duration>4</duration>
    <voice>3</voice>
    <type size="cue">eighth</type>
    <accidental parentheses="yes">natural</accidental>
    <stem default-y="-67">down</stem>
    <staff>2</staff>
    <beam number="1">continue</beam>
  </note>
  <note default-x="284">
    <pitch>
      <step>G</step>
      <octave>3</octave>
    </pitch>
    <duration>4</duration>
    <voice>3</voice>
    <type size="cue">eighth</type>
    <stem default-y="-67">down</stem>
    <staff>2</staff>
    <beam number="1">end</beam>
  </note>
</measure>
<!--=====-->

```

## PRIEDAS NR.4 DUOMENŲ APDOROJIMO PROGRAMINIAI KODAI

### Failas vertikal.sas

```
PROC IMPORT OUT= v.chrom
      DATAFILE= "&path"
      DBMS=EXCEL REPLACE;
      SHEET="Sheet1";
      GETNAMES=YES;
      MIXED=NO;
      SCANTEXT=YES;
      USEDATE=YES;
      SCANTIME=YES;

RUN;

proc sort data=v.chrom;
by grupe;
run;

proc freq data=v.chrom;
title "Pasiskirstymas 3/4";
table b/out=v.dazniai nocum;
by grupe;
run;

data v.vert_daz;
set v.dazniai;
length _01-_12 8.;
if substr(b,1,1)="1" then _01=1;
if substr(b,2,1)="1" then _02=1;
if substr(b,3,1)="1" then _03=1;
if substr(b,4,1)="1" then _04=1;
if substr(b,5,1)="1" then _05=1;
if substr(b,6,1)="1" then _06=1;
if substr(b,7,1)="1" then _07=1;
if substr(b,8,1)="1" then _08=1;
if substr(b,9,1)="1" then _09=1;
if substr(b,10,1)="1" then _10=1;
if substr(b,11,1)="1" then _11=1;
if substr(b,12,1)="1" then _12=1;
run;

proc freq data=v.vert_daz;
table _01/out=v.daz nocum;
by grupe;
weight _01;
run;

proc summary data=v.vert_daz;
var _01-_12/ weight=percent;
by grupe;
output out=v.vertikali (drop=_type_ _freq_) sumwgt=_01-_12;
run;
```

## Failas vert\_xml.sas

```
filename kurin "&path";
filename map 'C:\Users\HT\Desktop\RITMO ANALIZE\Kodai\import.map';
libname kurin xml xmlmap=map;
data v.kurinys;
set kurin.natos;
run;
data v.ritmas;
set v.kurinys;
tie=lag(tie);
if slur=lag(slur)+2 & step=lag(step) then slur1=0;
format taktas 4. tie 4. slur 4.;
slur=slur1;
drop slur1;
run;

* Keiciam tie = 0 jei nata uzlyguota (t.y. negrojama).
pataisomi lygavimai - tie ir slur, tada paleidziamas 32nd_paruosti --> 16th_taktai;
*kopija;
data v.ritmas1;
set v.ritmas;
run;
data v.trioles;
set v.ritmas;
if actual_notes^=.;
yra=1;
keep taktas yra;
run;
data v.trioles;
merge v.trioles v.ritmas;
by taktas;
run;
data v.ritmas;
set v.trioles;
if yra=.;
drop yra actual_notes normal_notes;
run;
data v.ritmas;
set v.ritmas;
format dot 3.;
if dot=. then dot=0;
if dot=lag(dot)+1 then kiek=1;
if dot=lag(dot)+2 then kiek=2;
*if _N_=1 & dot^=0 then kiek=dot;
run;
data v.ritmas;
set v.ritmas;
if chord^=lag(chord) then delete;
drop chord;
run;
data v.ritmas;
set v.ritmas;
format grace 3.;
```

```

if grace^=lag(grace) then delete;
drop grace;
run;
*-----ismetami tusti taktai-----;
data v.tvarka;
set v.ritmas;
if type^="";
if not(type="whole" and step="");
drop slur dot;
run;
* 64th taktu ismetimas;
data v.t64th;
set v.tvarka;
if type="64th" then yra=1;
keep taktas yra;
run;
data v.t64th;
if yra^=.;
merge v.t64th v.tvarka;
by taktas;
run;
data v.taktai_su64th;
set v.t64th;
if yra=1;
drop yra;
run;
data v.tvarka;
if yra^=. then do;
    set v.t64th;
    if yra=.;
    drop yra;
end;
else set v.tvarka;
run;
*-----priskiriami numeriai-----;
data v.tvarka;
set v.tvarka;
if type="32nd" then trukme=1;
if type="16th" then trukme=2;
if type="eighth" then trukme=4;
if type="quarter" then trukme=8;
if type="half" then trukme=16;
if type="whole" then trukme=32;
if kiek=1 then trukme=trukme*1.5;
if kiek=2 then trukme=trukme*1.75;
drop kiek;
run;
data v.tvarka;
set v.tvarka;
suma+trukme;
t2=mod(suma,32*metras/4);
if trukme=32*metras/4 then t2=0;
if _N_=1 then takt=1;

```

```

retain;
if lag(t2)=0 then takt+1;
run;
*--jei yra klaidingu taktu;
data v.klaidos;
run;
data v.klaidos;
set v.tvarka;
if taktas=lag(taktas) then summ+trukme;
else summ=trukme;
sum2=lag(summ);
run;
data v.klaidos;
set v.klaidos;
if taktas^=lag(taktas);
if (sum2^=(32/4*metras) & sum2^=(32/4*metras*2) &
sum2^=(32/4*metras*3) & sum2^=(32/4*metras*4)) then klaida=1;
keep taktas sum2;
if klaida=1;
taktas=taktas-1;
if taktas=0 then delete;
run;
data v.tvarkal;
merge v.tvarka v.klaidos;
by taktas;
suma=sum2;
drop sum2 t2 takt;
run;
data v.tvarka;
set v.tvarkal;
if suma=.;
drop suma;
run;
data v.tvarka;
set v.tvarka;
suma+trukme;
t2=mod(suma,32*metras/4);
if trukme=32*metras/4 then t2=0;
if _N_=1 then takt=1;
retain;
if lag(t2)=0 then takt+1;
run;

*----sujungiama----;
data v.tvarka;
set v.tvarka;
a=round(10**(trukme-1),1);
length b $32.;
b=substrn(a,1,32);
b=strip(b);
if trukme=32 then b="10000000000000000000000000000000";
if trukme=28 then b="100000000000000000000000000000";
if trukme=24 then b="1000000000000000000000000000";

```

```

if trukme=16 then b="1000000000000000";
if trukme=14 then b="1000000000000000";
if step="" or tie=0 then substr(b,1,1)="0";
drop a;
run;
*-----transponuoja ir sujungia-----;
proc transpose data=v.tvarka out=v.trans(drop=_name_);
var b;
by takt;
run;
data v.trans;
set v.trans;
length a $32.;
a=catt(of col:);
keep takt a;
run;
data v.metras;
set v.tvarka;
if takt=lag(takt) & metras=lag(metras) then delete;
keep takt metras;
run;
data v.trans;
merge v.trans v.metras;
by takt;
run;
*-----sesioliktiniu tikslumu-----;
data v.trans16;
set v.trans;
tr=0;
do i=1 to 16;
  k=substr(a,2*i,1);
  if k="1" then tr=1;
end;
do i=1 to 16;
  if tr=0 then do;
    substr(a,2*i,1)='2';
    b=compress(a,'2');
  end;
  else b1=a;
end;
drop k a i tr;
run;
data v.analizei;
set v.trans16;
if b1="";
drop b1;
grupe="papild_kurinys";
run;
data v.taktai32nd;
set v.trans16;
if b1^="";
drop b;
run;

```

```

quit;
PROC EXPORT DATA= v.ANALIZEI
            DBMS=EXCEL REPLACE
            OUTFILE= "&exelio\papildomas.xls";
            SHEET='Sheet1';
RUN;

```

### Failas vert\_paruosta.sas

```

PROC IMPORT OUT= v.vertikali
            DATAFILE= "&path"
            DBMS=EXCEL REPLACE;

SHEET="Sheet1";
GETNAMES=YES;
MIXED=NO;
SCANTEXT=YES;
USEDATE=YES;
SCANTIME=YES;

RUN;

```

### Failas pasiskirst.sas

```

pattern1 color=grey;
pattern2 color=very_light_blue;
pattern3 color=pink;
pattern4 color=light_green;
pattern5 color=light_blue;
pattern6 color=light_red;
pattern7 color=grey;
pattern8 color=light_orange;
pattern9 color=grayish_purple;

proc sort data=v.vertikali;
by grupe;
run;

proc transpose data=v.vertikali out=v.trans2;
var _01-_12;
by grupe;
run;

data v.trans2;
set v.trans2;
length kirtis 8.;
kirtis=substr(_name_,2,2);
gr=substr(grupe,1,3);
if grupe="ciurlionio_ankstyv" then gr="a_c";
if grupe="ciurlionio_velyv" then gr="v_c";
rename coll=daznis;

run;

filename fout1 "&rez\palyginimas.gif";
filename fout3 "&rez\palyginimas2.gif";
filename fout2 "&rez\histograma.gif";
ods html file ="&rez\pasiskirstymas.html";
goptions gsfname=fout1 device=gif gsfmode=replace;
title 'Pasiskirstymas pagal takto kircius';

```

```

proc gchart data=v.trans2;
vbar gr/ sumvar=daznis discrete group=kirtis patternid=group;
label gr="grupe";
run;
goptions gsfname=fout3 device=gif gsfmode=replace;
title 'Pasiskirstymas pagal grupes';
proc gchart data=v.trans2;
vbar kirtis/ sumvar=daznis discrete group=grupe patternid=group;
run;
goptions gsfname=fout2 device=gif gsfmode=replace;
proc univariate data=v.trans2;
var kirtis;
freq daznis;
title "Takto kirciu pasiskirstymas";
histogram kirtis/ midpoints=(1 2 3 4 5 6 7 8 9 10 11 12)
vaxislabel="Tankis, %";
run;
ods html close;
title;
quit;

```

### Failas exp.sas

```

PROC EXPORT DATA= v.vertikali
            OUTFILE= "&path"
            DBMS=EXCEL REPLACE;
            SHEET="exp";
RUN;

```



## PRIEDAS NR.5 MODELIŲ PROGRAMINIAI KODAI

### Failas corresp.sas

```
filename fout "&rez\atitikties.gif";
ods html file="&rez\atitikties_analize.html";
%let plotitop = gopts = gsfmode = replace
                device=gif
                gsfname=fout;

title "Paprastasis atitikties tyrimas";
data v.transp;
set v.vertikali;
n=cat("col",_N_);
run;
proc transpose data=v.transp out=v.transp(drop=_label_);
id n;
idlabel grupe;
run;
proc corresp data=v.transp out=v.corresp;
var col;
id _name_;
run;
%plotit(data=v.corresp, datatype=corresp, plotvars=Dim1 Dim2, href=0, vref=0)
ods html close;
```

### Failas dispersine.sas

```
filename fout1 "&rez\disp_prielaidos1.gif";
filename fout2 "&rez\disp_prielaidos2.gif";
ods html file="&rez\dispersine_grup.html";
proc glm data=v.trans2;
class grupe;
model kirtis=grupe;
freq daznis;
output out=v.paklaidos p=yhat student=pakl;
means grupe/ tukey lines;
run;
goptions gsfname=fout1 device=gif gsfmode=replace;
axis1 label=('Standartizuotos paklaidos');
axis2 label=("Prognozuojamos reikšmės");
symbol1 value=dot color=green;
title;
proc gplot data=v.paklaidos;
plot pakl*yhat/vaxis=axis1 haxis=axis2;
run;
goptions gsfname=fout2 device=gif gsfmode=replace;
proc gplot data=v.paklaidos;
plot pakl*grupe/vaxis=axis1;
run;
proc univariate normal;
var pakl;
run;
```

```
ods html close;
quit;
```

### Failas klast\_grup.sas

```
filename fout1 "&rez\klaster1.gif";
filename fout2 "&rez\klaster2.gif";
proc distance data=v.vertikali method=Euclid out=v.pan1;
var ratio(_01--_12);
id grupe;
run;
proc distance data=v.vertikali method=dcorr out=v.pan2;
var ratio(_01--_12);
id grupe;
run;
ods html file="&rez\klasterine_grup.html";
proc cluster data=v.pan1 method=&met outtree=v.Tree pseudo;
id grupe;
run;
goptions gsfname=fout1 device=gif gsfmode=replace;
title 'Grupavimas pagal Euklido atstuma';
proc tree data=v.tree horizontal;
id grupe;
run;
proc cluster data=v.pan2 method=&met outtree=v.Tree2 pseudo;
id grupe;
run;
goptions gsfname=fout2 device=gif gsfmode=replace;
proc tree data=v.tree2 horizontal;
id grupe;
title 'Grupavimas pagal panasumo koeficienta';
run;
ods html close;
quit;
```

### Failas klast\_tak.sas

```
proc transpose data=v.vertikali out=v.trans;
run;
data v.trans;
set v.trans;
label _name_='takto kirtis';
rename _name_=takt;
drop _label_;
run;
filename fout1 "&rez\klaster3.gif";
filename fout2 "&rez\klaster4.gif";
%let n1=Euclid;
%let n2=dcorr;
proc distance data=v.trans method=&n1 out=v.pan1;
var ratio(col:);
id takt;
run;
```

```

proc distance data=v.trans method=&n2 out=v.pan2;
var ratio(col:);
id takt;
run;
ods html file="%rez\klasterine_taktu.html";
proc cluster data=v.pan1 method=&met outtree=v.Tree pseudo;
id takt;
run;
goptions gsfname=fout1 device=gif gsfmode=replace;
title "Grupavimas. Metodas: &n1.";
proc tree data=v.tree horizontal;
id takt;
run;
proc cluster data=v.pan2 method=&met outtree=v.Tree2 pseudo;
id takt;
run;
goptions gsfname=fout2 device=gif gsfmode=replace;
proc tree data=v.tree2 horizontal;
id takt;
title "Grupavimas. Metodas: &n2.";
run;
ods html close;
quit;

```

### Failas mca.sas

```

filename fout "%rez\mca.gif";
ods html file="%rez\mca_rezultatai.html";
%let plotitop = gopts = gsfmode = replace
                device=gif
                gsfname=fout;
title "Daugialype atitikciu analize";
proc corresp data=v.trans2 mca outc=v.rez2;
tables grupė kirtis;
weight daznis;
run;
%plotit(data=v.rez2, datatype=corresp, href=0, vref=0)
ods html close;

```

### Failas npar.sas

```

ods html file="%rez\npar_grup.html";
proc nparlway data=v.trans2 wilcoxon median vw savage;
class grupė;
var kirtis;
freq daznis;
run;
ods html close;

```

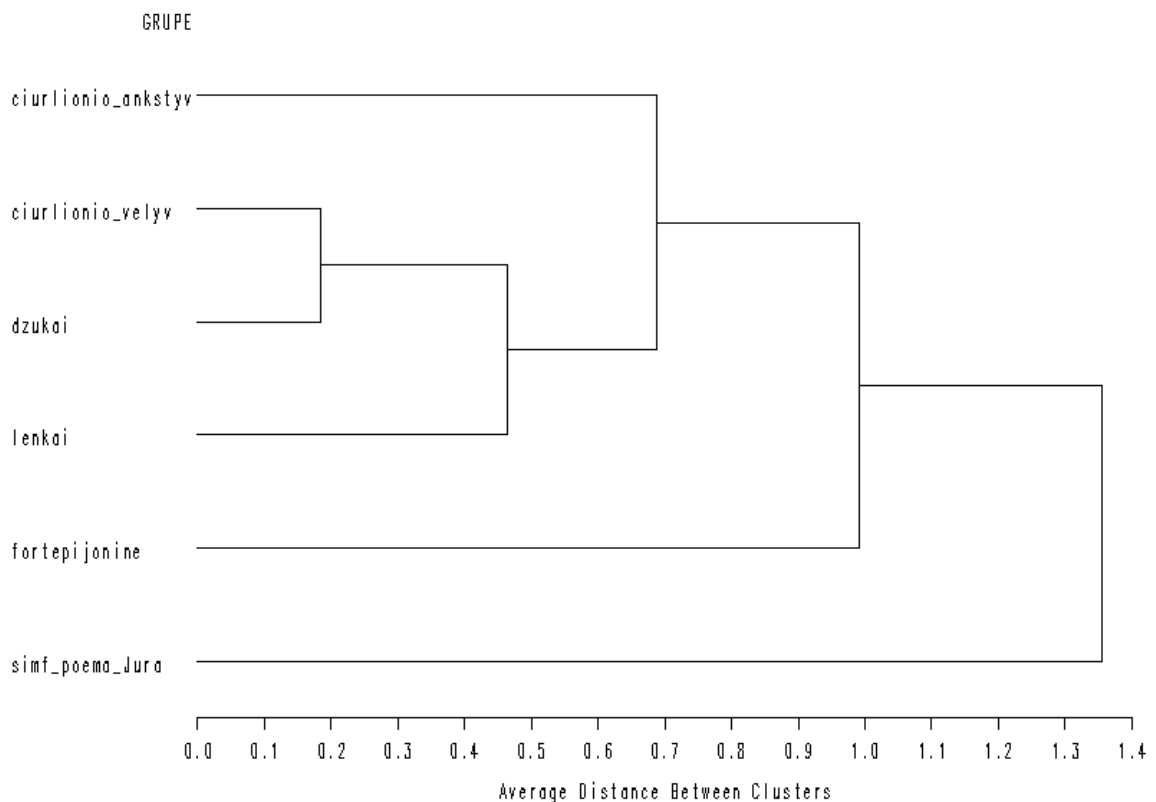
## Failas pan\_grup.sas

```
proc distance data=v.vertikali method=Euclid out=v.pan1;
var ratio(_01--_12);
id grupe;
run;
proc distance data=v.vertikali method=shape out=v.pan11;
var ratio(_01--_12);
id grupe;
run;
proc distance data=v.vertikali method=simratio out=v.pan2;
var ratio(_01--_12);
id grupe;
run;
proc distance data=v.vertikali method=corr out=v.pan22;
var ratio(_01--_12);
id grupe;
run;
ods html file="%rez\panasumo_koef.html";
proc print data=v.pan1;
title 'Euklido atstumai';
proc print data=v.pan11;
title 'Formos atstumai';
proc print data=v.pan2;
title 'Panašumo koeficientai';
proc print data=v.pan22;
title 'Koreliacijos panašumo koeficientai';
run;
ods html close;
```

## PRIEDAS NR.6 TYRIMO REZULTATŲ PAVYZDŽIAI

### 1 pavyzdys.

Pateikiami visų analizuotų tyrime kūrinių grupių svarbiausi rezultatai, kurie nebuvo pateikti tiriamojoje dalyje. Jei lyginsime gautus klasterizavimo rezultatus tik pagal atstumus, nepriklausomai nuo to koks metodas taikytas klasteriams apjungti, pastebėsime, kad išsiskiria vienintelio koreliacijos koefeciento paversto Euklido atstumu grupavimo medis. Šiuo atveju nėra atskiriamos dvi grupės, o prie pradinio klasterio kiekvienam etape prijungiama po vieną kūrinių grupę (žr. Pav.1).



**Pav. 1 Kūrinių grupių klasterizavimas pagal *dcorr* atstumą**

Toliau pateikti atitikties analizės rezultatai eilučių taškams, kur kategoriniai kintamieji – taktų kirčiai. Lentelėje Nr.1 pateiktos eilučių taškų koordinatės, lentelėje Nr.2 – abiejų koordinatinių kokybė, eilučių masės visos dažnių lentelės atžvilgiu ir proporcija arba dalis kurią sudaro kiekvienas taškas visos bendrosios inercijos atžvilgiu.

**Lentelė Nr. 1 Takto kirčių - eilučių taškų koordinatės**

Takto kirtis	Dim1	Dim2
_01	0,2240	-0,0269
_02	-0,7859	0,4641

Takto kirtis	Dim1	Dim2
_03	-0,2324	-0,2264
_04	0,2604	0,6978
_05	0,2116	0,0216
_06	-1,0888	0,3663
_07	-0,4461	-0,1474
_08	-0,9258	0,2505
_09	-0,0394	-0,1742
_10	-0,4195	0,5090
_11	0,0102	0,1307
_12	-0,1262	0,4693

**Lentelė Nr. 2 Eilučių taškų aprašomosios statistikos**

Takto kirtis	Kokybė	Masė	Inercija
_01	0,9852	0,2472	0,0895
_02	0,8667	0,0120	0,0805
_03	0,8431	0,0963	0,0842
_04	0,9503	0,0300	0,1227
_05	0,8430	0,2004	0,0753
_06	0,9365	0,0099	0,0973
_07	0,9211	0,0773	0,1297
_08	0,8092	0,0146	0,1164
_09	0,8614	0,1718	0,0445
_10	0,7950	0,0158	0,0607
_11	0,2967	0,1032	0,0418
_12	0,6244	0,0216	0,0573

Matome, kad visų taškų kokybė pirmų dviejų dimensijų atžvilgiu yra pakankamai didelė ir viršija 0,8, išskyrus 10, 11, 12 taškus. Dimensijų atžvilgiu stulpelių taškų koordinatų patikimumas gali būti interpretuojamas pagal lentelės Nr.3 rezultatus, arba lentelės Nr.4

indeksus. Geriausiai apibūdinančios tašką dimensijos indeksas parašomas paskutiniame stulpelyje. Šiuo atveju 1,2, 5 – 8 takto kirčius geriausiai apibūdina pirmoji dimensija, sudaranti 55,34% visos inercijos, likusius takto kirčius geriau apibūdina antroji dimensija, sudaranti 29,53% visos inercijos. 19,47% pirmosios dimensijos inercijos sudaro 7 takto kirtis, taip pat virš 10% sudaro kiekvienas takto kirtis iš išvardintų – 1, 5, 6 ir 8.

**Lentelė Nr. 3 Eilučių taškų inercija dimensijų atžvilgiu**

Takto kirtis	Dim1	Dim2
_01	0.1570	0.0042
_02	0.0935	0.0611
_03	0.0658	0.1171
_04	0.0258	0.3466
_05	0.1136	0.0022
_06	0.1479	0.0314
_07	0.1947	0.0399
_08	0.1586	0.0218
_09	0.0034	0.1236
_10	0.0353	0.0973
_11	0.0001	0.0418
_12	0.0044	0.1129

**Lentelė Nr. 4 Eilučių taškų koordinatinių indeksai pagal inerciją**

Takto kirtis	Dim1	Dim2	Geriausias
_01	1	0	1
_02	1	1	1
_03	0	2	2
_04	0	2	2
_05	1	0	1
_06	1	0	1
_07	1	0	1
_08	1	0	1

Takto kirtis	Dim1	Dim2	Geriausias
_09	0	2	2
_10	0	2	2
_11	0	0	2
_12	0	2	2

Remdamiesi atitikties žemėlapiu pateiktu paveiksle 2.7, galime įvertinti 1 ir 5 takto kirčių artumą, taip pat pirmosios dimensijos ir atstumo tarp taškų atžvilgiu galėtume sugrupuoti 3 ir 7 takto kirčius, bei 2, 6 ir 8.

### 2 pavyzdys.

Pateikiamas palyginamosios analizės tyrimo pavyzdys, kai lyginamas papildomas kūrinys su visomis tirtomis kūrinų grupėmis. Galime pastebėti, kad papildomo kūrinio histograma yra tarp fortepijoninės ir lenkų liaudies muzikos formos atžvilgiu (žr. Pav.2). Žemiau pateikti dispersinės analizės rezultatai, hipotezė apie vidurkių lygybę atmesta, kategorijos yra reikšmingos taktų kirčių atžvilgiu (žr. lentelę Nr.5). Tačiau apibrėžtumo koeficientas (R-kvadaratu) rodo, kad modelio rezultatai apibūdina tik 1,28% variacijos.

**Lentelė Nr. 5 Dispersinės analizės rezultatai**

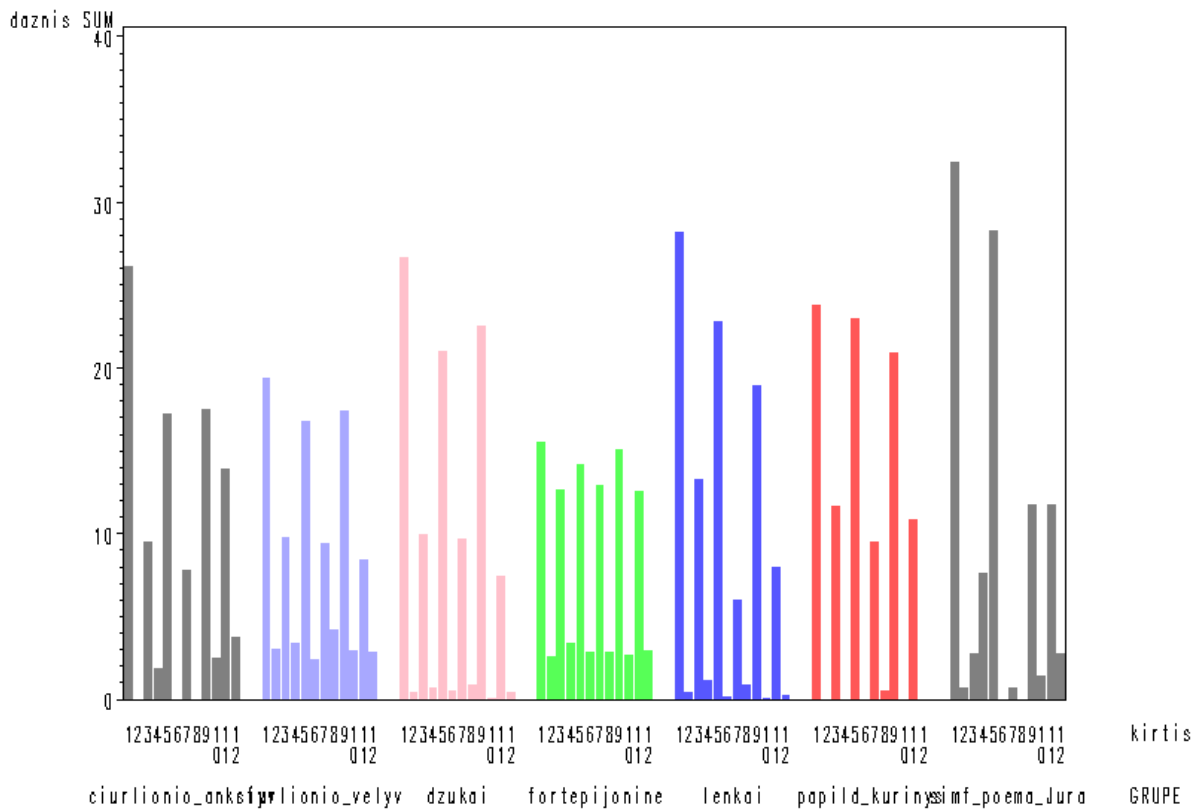
Nuokrypų šaltinis	LL	Nuokrypių kvadratų sumos	Nuokr. kvadratų vidurkiai	Fišerio statistika	p-reikšmė
Faktorius <i>grupe</i>	6	408,7763	68,12938	5,55	<,0001
Atsitiktinis faktorius	2564	31476,12	12,27618		
Visi faktoriai	2570	31884,9			
R-kvadratas	Variacijos koeficientas	Šaknis iš MSE	Dydžio kirtis vidurkis		
0,012820	63,25920	3,503738	5,538701		



**Lentelė Nr. 6 Tjukio grupavimas kūrinių grupėms su papildomu kūriniu**

Tjukio grupavimas	Vidurkis	N	Grupė	
A	6.0692	477	fortepijonine	
A				
A	5.8035	341	ciurlionio_ankstytv	
A				
A	5.7943	423	ciurlionio_velyv	
A				
B	A	5.5014	349	papild_kuriny
B	A			
B	A	5.3035	369	dzukai
B				
B		4.9641	334	lenkai
B				
B		4.9640	278	simf_poema_Jura

Šiuo atveju Tjukio grupavimas pateiktas lentelėje Nr.6 pateikia du grupavimo būdus: A – sugrupuojama visa nagrinėjama Čiurlionio kūryba su dzūkų liaudies muzikos kūriniais, B – visa nagrinėjama liaudies muzika su simfonine pema „Jūra“ ir papildomu kūriniu. Tačiau dispersinės analizės rezultatai nėra patikimi, nes netenkinamos prielaidos ir labai maža apibrėžtumo koeficiento reikšmė.



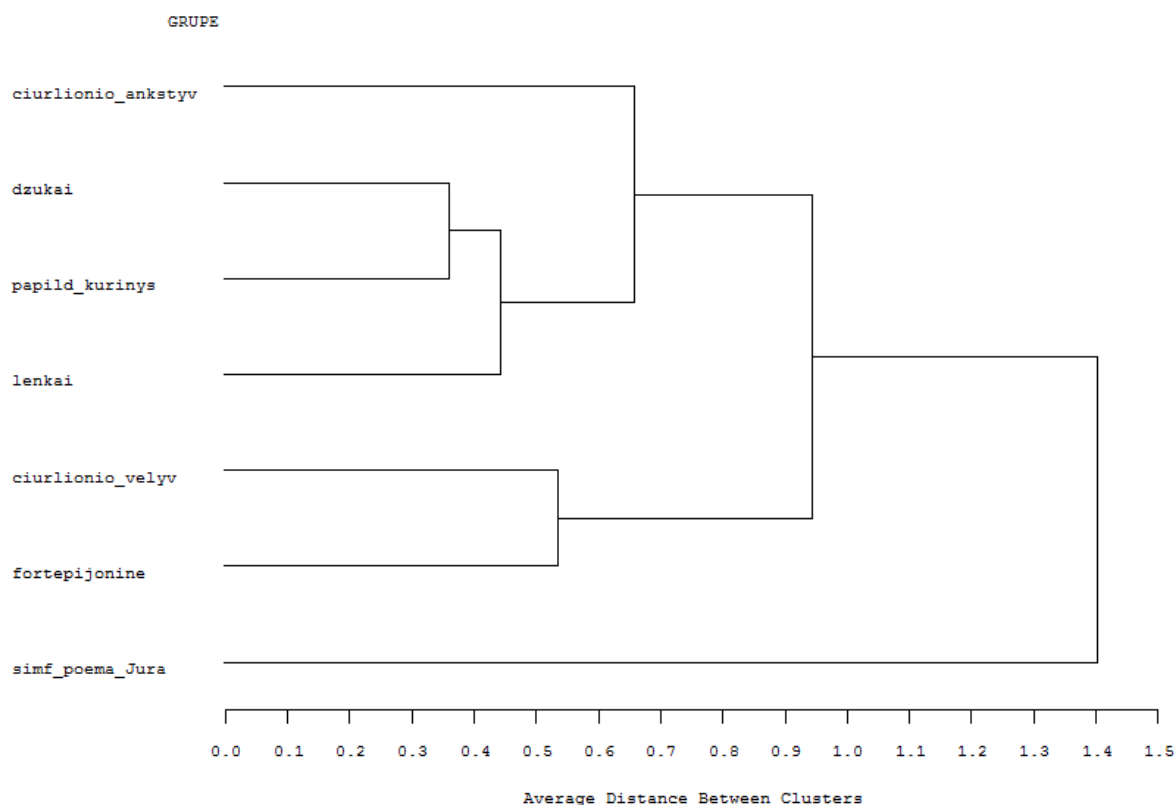
**Pav. Nr. 2 Pasiskirstymas pagal grupes, palyginimui su papildomu kūrinium**

Neparametrinė dispersinė analizė rodo analogiškus rezultatus, hipotezė apie vidurkių lygybę kategorijų atžvilgiu priimama su reikšmingumo lygmeniu 0,05, visų kriterijų atveju (žr. lentelę Nr.7).

**Lentelė Nr. 7 Neparametrinių kriterijų rezultatai**

Kriterijus	Chi-kvadrato statistika	p-reiškė
Wilcoxon	8,5110	0,2030
Medianos	10,8258	0,0939
Van der Waerden	8,9076	0,1788
„Savage“	8,7962	0,1854

Žinome, kad papildomas kūrinys pasirinktas iš fortepijoninės muzikos (nors į analizuojamos fortepijoninės muzikos rinkinį neįeina), tačiau visų naudojamų koeficientų atžvilgiu, kūrinys yra artimesnis dzūkų ir lenkų liaudies muzikai, bei ankstyvajai Čiurlionio kūrybai. Tai matome ir žemiau pateiktame grupavimo medyje.

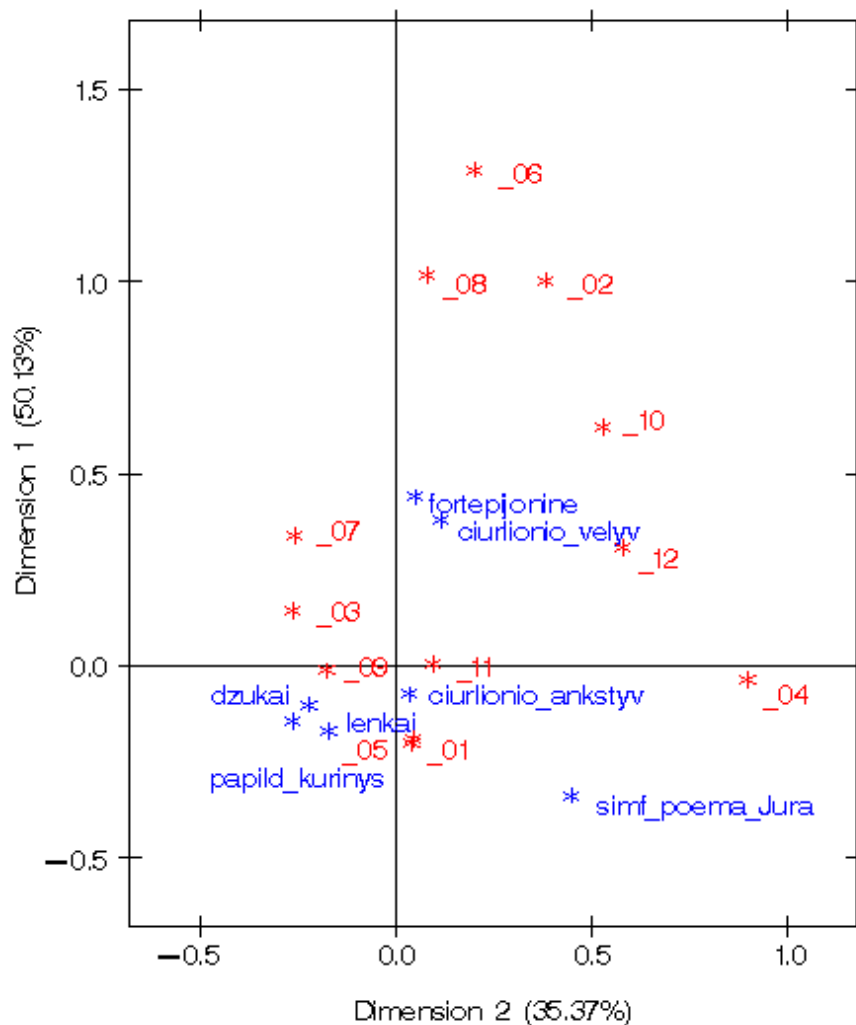


**Pav. Nr. 3 Kūrinių grupių klasterių dendrograma**

Atitikties analizės žemėlapis (žr. pav.Nr.4) rezultatai rodo, kad papildomas kūrinys grupuojamas su lenkų ir dzūkų liaudies muzika. Koordinačių kokybė pirmųjų dviejų dimensijų atžvilgiu yra didelė (viršija 0.75), tačiau šių taškų inercija maža lyginant su kitų grupių taškų inercija. Šiuo atveju Čiurlionio vėlyvoji ir fortepijoninė muzika yra labai artimos pirmosios dimensijos atžvilgiu, be to jų koordinačių kokybė yra didelė ir jų taškai sudaro didelę dalį visos inercijos. Simfoninė poema „Jūra“ turi didelę inerciją ir didelę koordinačių kokybę, tačiau neturi artimų eilučių taškų.

**Lentelė Nr.8 Kūrinių grupių taškų aprašomosios statistikos**

Grupė	Kokybė	Masė	Inercija
ciurlionio_ankstytv	0,0842	0,1429	0,0772
ciurlionio_velyv	0,8824	0,1429	0,1739
dzukai	0,8175	0,1429	0,0687
fortepijonine	0,9389	0,1429	0,2032
lenkai	0,7566	0,1429	0,0731
papild_kurinys	0,9284	0,1429	0,0935
simf_poema_Jura	0,9859	0,1429	0,3103



**Pav. Nr.4 Atitikties žemėlapis palyginimui su papildomu kūriniumi**

Palyginsime gautą grupavimą su J-divergencijos gautais rezultatais lentelėje Nr.9. Stipresniųjų takto dalių atžvilgiu apskaičiuota Jeffreys simetrinė divergencija rodo, kad artimiausi papildomam kūriniumi iš lyginamų objektų yra Čiurlionio vėlyvoji ir dzūkų liaudies muzika.

**Lentelė Nr.9 J-divergencija papildomam kūriniumi su kūriniumi grupėmis**

<i>Papildomo kūrinio Jeffreys divergencija</i>	
Čiurlionio ankstyvoji	0.0513
Čiurlionio vėlyvoji	0.0103
Dzūkų liaudies	0.0312
Čiurlionio fortepijoninė	0.1268
Lenkų liaudies	0.0520
Simf. poema „Jūra“	0.6592

## PRIEDAS NR.7 DARBE NAUDOJAMI MUZIKINIAI TERMINAI

**Nata** – garsą vienoje oktavoje apibūdinantis simbolis, vienas iš 7 simbolių A-G. Ritmo atžvilgiu nata reiškia garso buvimą.

**Pauzė** – muzikoje apibūdina tylą. Šis dydis turi ilgį taip pat kaip ir nata.

**Natos ilgis** – natos ar pauzės trukmę apibūdinantis parametras. Įgyjamos reikšmės – mažėjimo tvarka dalinant kiekvieną reikšmę pusiau: pilnoji, pusinė, ketvirtinė, aštuntinė, šešioliktainė, trisdešimtantrinė, šešiasdešimtketvirtinė.

**Papuošimas** (angl.*grace*) – trumpoji nata, dažniausiai muzikiniame tekste būna sumažinta ir rašoma prieš norimą natą kaip muzikinis papuošimas. Ritmo atžvilgiu ši nata neturi ilgio.

**Triolės** ir kitos nestandartinio ilgio natos (angl.*tuplet*) – natų, neatitinkančių tikrosios jų trukmės grupės, pavyzdžiui triolė yra sudaryta iš trijų aštuntinių, kurios takte užima dviejų aštuntinių natų trukmę.

**Taškas** – simbolis reiškiantis natos ilgio pratęsimą pusę jos vertės.

**Apjungimas** – dviejų natų sujungimas lanku, reiškiančių pirmosios natos tęstinumą arba tolygų perėjimą tarp natų.

**Akordas** – apibūdina kelių natų, skambančių vienu metu, grupę.