

MYKOLO ROMERIO UNIVERSITETAS

Vitalija Rudzkienė

SOCIALINĖ STATISTIKA

Vadovėlis

Vilnius 2005

UDK 313:316(075.8)

Ru44

2004 m. gruodžio 10 d. Nr. A-241

Lietuvos Respublikos švietimo ir mokslo ministerijos Aukštųjų mokyklų bendrųjų vadovėlių leidybos komisijos rekomenduota

Recenzavo:

Vilniaus Gedimino technikos universiteto Matematinės statistikos katedros vedėjas prof. habil. dr. **Leonas Saulis**;

Mykolo Romerio universiteto Žinių visuomenės instituto direktorius prof. dr. **Arūnas Augustinaitis**;

Vilniaus universiteto Kiekybinių metodų ir modeliavimo katedros doc. dr. **Vytautas Kėdaitis**

Vadovėlis svarstytas Lietuvos teisės universiteto Valstybinio valdymo fakulteto Teisinės informatikos katedros 2004 m. gegužės 26 d. posėdyje (protokolo išrašas Nr. 1TIK-4) ir rekomenduotas spausdinti

Lietuvos teisės universiteto vadovėlių, monografijų, mokslinių, mokymųjų, metodinių bei kitų leidinių aprobavimo spaudai komisija 2004 m. gegužės 26 d. posėdyje (protokolas Nr. 2L-5) vadovėlį patvirtino spausdinti

Išleista parėmus Lietuvos Respublikos švietimo ir mokslo ministerijai

Visos leidinio leidybos teisės saugomos. Šis leidinys arba kuri nors jo dalis negali būti dauginami, taisomi ar kitu būdu platinami be leidėjo sutikimo.

ISBN 9955-19-002-7

© Mykolo Romerio universitetas, 2005

© Vitalija Rudzkienė, 2005

TURINYS

Pratarmė.....	6
Socialinės statistikos objektas ir uždaviniai.....	9
Socialinių žinių struktūra: duomenys, modeliai ir teorijos.....	13
1. Socialiniai tyrimai ir viešoji nuomonė.....	18
1.1. Apibrėžimai.....	19
1.2. Statistinės informacijos rinkimo būdai.....	21
1.3. Klausimų formulavimas ir apklausos lapų sudarymas.....	24
1.4. Tyrimų rezultatų analizė ir ataskaita.....	26
1.5. Imčių sudarymo būdai.....	28
1.6. Matavimų skalės.....	30
1.7. Imties dydis.....	32
1.8. Apklausos taisyklės, apklausos anketų pavyzdžiai.....	35
Kontroliniai klausimai.....	45
2. Pirminė duomenų apžvalga arba aprašomoji statistika.....	46
2.1. Dažnių analizė.....	47
2.2. Grafinis duomenų vaizdavimas.....	52
2.3. Duomenų padėties ir sklaidos charakteristikos.....	54
2.4. Nesudėtingos apklausos anketos pradinės duomenų analizės pavyzdys.....	60
Užduotys.....	81
3. Demografinė statistika.....	83
3.1. Demografinių duomenų šaltiniai.....	84
3.2. Demografiniai rodikliai ir grupės.....	88
3.3. Struktūros dalių santykiniai dydžiai.....	91
3.4. Demografinių duomenų grafinis vaizdavimas.....	93
3.5. Demografinių duomenų pradinės analizės pavyzdys.....	96
Užduotys.....	99

3.6. Demografinės prognozės modeliai.....	102
3.7. Laiko eilutės ir aproksimacijos metodai.....	104
3.8. Aproksimacijos metodo taikymo pavyzdys.....	108
Užduotys.....	111
3.9. Kohortų (grupių) komponentiniai metodai.....	112
3.10. Tikėtina gyvenimo trukmė ir amžiaus lentelės.....	115
Užduotys.....	119
4. Statistinė gyvenimo lygio analizė: visuotinė gerovė ir nelygybė.....	122
4.1. Skurdo sąvoka ir rodikliai.....	124
4.2. Gyventojų pajamų ir išlaidų tyrimai.....	128
4.3. Pajamų pasiskirstymo netolygumo vertinimo metodai.....	135
4.4. Pajamų pasiskirstymo netolygumo įvertinimo pavyzdys.....	139
Užduotys.....	144
5. Užimtumo ir nedarbo statistika.....	148
5.1. Bedarbiai ir nedarbo tipai.....	151
5.2. Nedarbas ir skurdas.....	152
5.3. Statistinės užimtumo ir nedarbo sąvokos ir rodikliai.....	153
5.4. Pagrindinių užimtumo ir nedarbo rodiklių skaičiavimo pavyzdys.....	156
5.5. Struktūros skirtumų vertinimo pavyzdys.....	157
5.6. Gyventojų užimtumo ir nedarbo bei kitų socialinių rodiklių ryšių analizė.....	158
5.7. Pirsono χ^2 suderinamumo kriterijus.....	159
5.8. χ^2 kriterijaus skaičiavimas didesnėms nei 2×2 lentelėms.....	165
5.9. Dažnių lentelių ryšio stiprumo matai.....	167
Užduotys.....	168
6. Integruoti socialinės raidos rodikliai.....	171
6.1. Plačiausiai naudojami integruoti rodikliai.....	172
6.2. Socialinės pažangos indeksas.....	174
6.3. Gyventojų skurdo indeksas.....	176
6.4. Rodiklių konstravimo pagrindumas.....	177
6.5. Spirmeno koreliacijos koeficientas.....	178
6.6. Rodiklių skaičiavimo pavyzdžiai.....	182
Užduotys.....	185

7. Teisinė statistika.....	190
7.1. Teisiniai statistiniai rodikliai.....	190
7.2. Rodiklių tipai.....	193
7.3. Koreliacinė analizė. Pirsono koreliacijos koeficientas.....	195
7.4. Kriminalinių ir socialinių bei ekonominių rodiklių ryšiai ir jų analizės metodai.....	201
7.5. Tiesinė regresinė analizė.....	205
7.6. Tiesinė regresinė analizė <i>Statistica</i> programa.....	207
7.7. Regresinė analizė ir kokybiniai kintamieji.....	211
7.8. Netiesinė (kreivinė) regresija.....	219
Užduotys.....	222
8. Politinio gyvenimo statistika.....	227
8.1. Rinkimų prognozės.....	228
8.2. Politinio gyvenimo statistikos metodai.....	242
8.3. Logistinė regresija.....	245
8.4. Logistinė regresija <i>Statistica</i> programa.....	247
Užduotys.....	251
Literatūra.....	253
Priedai.....	256



PRATARMĖ

Šiuolaikinėje visuomenėje informacija, jos analizė ir socialinių bei ekonominių reiškinių prognozė tapo vienu iš svarbiausių strateginių uždavinių, leidžiančių suvokti sudėtingus pokyčius, kurie intensyviai vyksta ir vyks visoje planetoje, ir sudarančių galimybę kurti veiksmingas valdymo sistemas. Technologijų plėtra ir globali ekonomikos raida, be socialinio progreso koordinavimo, sukelia daugelį neigiamų pasekmių: visuomenės nuskurdinimą, sveikatos būklės blogėjimą, infekcinių ligų protrūkius, nevaldomą demografinę situaciją (visuomenės senėjimą ir migraciją), ekosistemų pokyčius ir gamtinių resursų išsekimą bei ilgalaikius klimato pokyčius. Demokratinės visuomenės raida pagrįsta egzistuojančios tikrovės pažinimu, kuris neįmanomas be statistinių bei socialinių metodų. Tikrovę galima pažinti tik stebint jos faktus bei procesus. Tuo tarpu kiekviena mokslo sritis, taip pat ir socialiniai mokslai, sukuria savo atskirą tikrovės modelį (jurisprudencijoje tai įstatymai, teisės aktai, nutarimai, teismų nutartys ir kt.). Šis abstraktus mąstymo sukurtas tikrovės modelis ne visada tiksliai atspindi egzistuojančią tikrovę ir net bando ją pakeisti sukurtais loginėmis konstrukcijomis. Egzistuojančiam pasauliui pažinti neužtenka vien loginio mąstymo. Empirinių stebėjimų gausa sukuria sąlygas dėsningumams, kitimo ypatybėms ir kitokioms žinioms, būtinoms šiuolaikinės visuomenės santykių tobulinimui, atsiskleisti. Statistikos metodai, paremti empiriniais stebėjimais, padeda suprasti ir numatyti visuomenėje vykstančius procesus, analizuoti jų paplitimą bei dinamiką ir gauti informacijos, reikalingos visuomenės santykiams tobulinti.

Dėl nuolatinio vis gausėjančio informacijos srauto, ypač per visuomenės informavimo priemones, socialinė statistika tampa priemone, lengvinančia skirtingų visuomenės sluoksnių tarpusavio supratimą ir susikalbėjimą. Socialiniai tyrimai yra mokslinis socialinių reiškinių suvokimo pagrindas. Per pastaruosius du dešimtmečius Lietuva, kaip ir daugelis kitų valstybių, patyrė išpūdingų socialinių, ekonominių ir technologinių pokyčių, ir šie pokyčiai vyksta toliau. Taigi dabar jau akivaizdu, kad pagrindinės ekonominės politikos pokyčių sukeltos problemos yra ne ekonominės, o socialinės. Socialinės politikos tobulinimas turi remtis

įvairių lygių socialinių sluoksnių, būtent šeimų, kaimynų, bendruomenių ir pan., charakteristikų ir sudėtingų jų tarpusavio ryšių supratimu. Šiai sudėtingai sistemai suvokti tarnauja socialiniai statistiniai tyrimai ir duomenų analizė.

Socialiniai duomenys gaunami dviem pagrindiniais būdais: gyventojų surašymu ir valstybės institucijų registruojama statistika (ekonomikos pokyčiai, sveikata, išsilavinimas, justicija, darbo jėga, gyventojų migracija ir pan.). Jeigu valstybinės statistinės institucijos netenkina socialinės informacijos poreikio, tai informacijos stygių kompensuoja kiti šaltiniai: sociologijos institutai, visuomenės nuomonės tyrimo centrai ir pan. Pasaulyje sukurtas platus socialinės, ekonominės ir teisinės statistinės informacijos tinklas. Nors statistinės informacijos srautas atrodo tiesiog neišsenkantis, tačiau gauti detalių mikroduomenų – vis dar sunkiai išsprendžiama problema. Kadangi efektyvūs daugiamačiai statistiniai metodai negali būti taikomi apibendrintiems duomenims, gauti pradinę informaciją yra būtina sąlyga.

Socialinės statistikos uždavinių sprendimas grindžiamas aprašomosios ir matematinės statistikos metodais – grafiniu duomenų vaizdavimu, dinamikos eilutėmis, duomenų padėties ir sklaidos charakteristikomis, koreliacija, hipotezių tikrinimu, regresine analize. Pokyčių socialinėje aplinkoje įvertinimo ir prognozavimo uždaviniuose labai svarbi griežta statistinių rezultatų interpretacija.

Skaitytojui pateikiamuose pavyzdžiuose dažniausiai nagrinėjami realūs duomenys ir, pasitelkiant šiuos pavyzdžius, aiškinami įvairių socialinių rodiklių skaičiavimo metodai bei šių metodų ir modelių taikymas socialiniams reiškiniams analizuoti ir prognozuoti. Uždaviniams spręsti naudojamos šiuolaikinės informacijos technologijos priemonės – kompiuteriai ir taikomoji programinė įranga. Statistinei analizei atlikti tinkamų priemonių yra ganėtinai daug. Galima naudoti *EXCEL*, *STATGRAPHICS*, *STATISTICA*, *SPSS*, *SAS* sistemas bei kitas informacijos technologijos priemones duomenų statistinei analizei atlikti. Pavyzdžiuose pateiktiems sudėtingesniems uždaviniams spręsti pasirinkta *STATISTICA* programa, kadangi daugelis universitetų turi šio programinio produkto licencijas.

Nuoširdžiai dėkoju visiems, kurie mane paskatino ir padėjo gilintis į sudėtingą, tačiau įdomų ir patrauklų socialinės statistikos pasaulį. Už skatinimą dėmesį socialiniais reiškiniais, atvirumą naujoms idėjoms, keliamus aukštus intelektualius reikalavimus, dėkoju savo kolegoms iš

Mykolo Romerio universiteto Valstybinio valdymo fakulteto. Šio universiteto kūrybinė atmosfera, kolegų pritarimas ir draugiška parama yra geriausias stimulas moksliniams tyrimams.

Dėkoju viešosios nuomonės ir rinkos tyrimų kompanijai „TNS GALLUP“ ir šios kompanijos kiekybinių tyrimų skyriaus vadovei Jurgai Verikienei už originalias apklausų anketas, kurios puikiai iliustruoja ir papildo vadovėlio medžiagą.



SOCIALINĖS STATISTIKOS OBJEKTAS IR UŽDAVINIAI

Terminas „socialinė statistika“ vartojamas dvejopai: kaip mokslo sritis ir kaip praktinė veikla.

Socialinę statistiką kaip *mokslo sritį* sudaro socialinių tyrimų teorija ir metodai. Tai skaitinės informacijos apie visuomenėje vykstančius socialinius reiškinius ir procesus rinkimo bei apibendrinimo metodai, kiekybiniai ir kokybiniai duomenų analizės ir prognozės metodai.

Socialinę statistiką kaip *praktinės veiklos sritį* reglamentuoja statistikos arba analogiškas įstatymas. Statistikos įstatymas numato valstybinės statistikos tikslus, duomenų gavimo būdus, duomenų rinkimo organizavimą, jų naudojimo sąlygas, taip pat atsakomybę už duomenų teikimo ir naudojimo tvarkos pažeidimus.

Šios abi sritys glaudžiai susijusios tarpusavyje. Surinktų statistinių duomenų nauda bei norimų gauti rezultatų tikslumas priklauso nuo teorinių prielaidų, kuriomis remiantis renkami duomenys ir nustatomi analizės metodai. Statistine sistema ar statistiniais metodais sudarytas socialinės tikrovės vaizdas panašus į fizinį kraštovaizdį žemėlapyje. Žemėlapis, kurį nori matyti geologas, labai skirsis nuo žemėlapio, kurio pageidautų taksi vairuotojas. Jei sudarant žemėlapių neaišku, kas juo naudosis, ir nėra aiškios sistemos, kurias kraštovaizdžio savybes reikia atskleisti, iš tokio žemėlapio didelės naudos nebus. Anot filosofo Wittgenstein, tai būtų panašu į tai, kad uždėjęs vieną ranką ant pakaušio pasakytum „Aš žinau savo ūgį“.

Socialinės statistikos sistemos, kaip ir žemėlapiai, naudojamos tam tikru tikslu, ir šis tikslas daugiau ar mažiau detalai bei suprantamai apibrėžia prielaidas ir principus, kuriais remiantis renkami duomenys. Jais remiantis parenkami ir analizės metodai.

Gal geriausia tiesiog numatyti tyrimo tikslus ir pagal juos konstruoti norimą statistinę sistemą? Tačiau dažniausiai bet kurios statistinės sistemos tyrimo tikslas – parodyti kitimo laike tendencijas, palyginti praeitį su dabartimi ir stengtis prognozuoti ateitį. Kadangi mes naudojame praeityje surinktus duomenis, tai kartu tenka naudotis ir šių duomenų

statistinėmis sistemomis. Gerai tai ar blogai, tačiau mes tarsi stovime duomenų ir informacijos upės viduryje ir dažnai kenčiame dėl mums naudingų žinių stygiaus ir nesuvokiamumo. Iš dalies priežastis yra ta, kad mes gyvename vadinamajame informacijos amžiuje. Ir nors pastaraisiais dešimtmečiais statistinių duomenų surenkama vis daugiau, į šią veiklą įsitraukė tarptautinės organizacijos (Jungtinės tautos, Pasaulio bankas ir kt.), pagrindiniai duomenys, kuriais matuojamas socialinis ir ekonominis progresas, mažai pakito. Tokio tipo pokyčiai vyksta palengva.

Kaip ir prieš daugelį metų, kiekviena šalis kasmet išleidžia statistinį metraštį. Statistinis metraštis yra pagrindinis valstybės statistinis leidinys. Jame pateikiami visos valstybės ir jos atskirų regionų geografiniai bei meteorologiniai duomenys, duomenys apie gyventojus, socialiniai rodikliai, gyventojų ekonominės veiklos rodikliai, kainos, gamyba ir paslaugos, finansai, nacionalinės sąskaitos ir nacionalinis turtas, kai kurių užsienio valstybių statistiniai duomenys.

Statistinis metraštis skiriamas platiesiems visuomenės sluoksniams. Be statistinių duomenų, jame pateikiamos svarbiausios sąvokos ir pagrindinių rodiklių skaičiavimo metodikos.

Statistinę informaciją gali rengti valstybinės ir nevalstybinės tarnybos. Statistinę informaciją skelbia visuomenės nuomonės tyrimo organizacijos, partijos, bankai ir kitos institucijos.

Viso pasaulio bei konkrečių kontinentų ir valstybių grupių demografinius ir socialinius bei ekonominius rodiklius renka ir skelbia tarptautinės organizacijos, pirmiausia Jungtinių Tautų (JT) organizacija. Pagrindinis šios organizacijos leidinys yra „Statistinis metraštis“ (*Statistical Yearbook*), leidžiamas anglų ir prancūzų kalbomis. Jungtinių Tautų organizacija ir jos specializuotos institucijos UNESCO ir FAO buvo įkurtos 1945 metais. JT kaupia ir skelbia duomenis apie visas savo nares, neatsižvelgiant į jų ekonominį išsivystymą. Kartu JT kuria statistinius standartus ir klasifikatorius tam, kad būtų galima palyginti atskirų valstybių pateikiamus duomenis.

Savo duomenis reguliariai publikuoja ir kitos tarptautinės organizacijos: ekonominio bendradarbiavimo ir plėtros organizacija (OECD), Europos Sąjungos statistikos valdyba (Eurostat), Pasaulio bankas (WB), Tarptautinis valiutos fondas (IMF) ir jo dukterinės institucijos bei įvairių regionų bankai, pavyzdžiui, Azijos plėtros bankas (ADB). OECD renka ir skelbia išsivysčiusių industrinių valstybių statistinius duomenis.

Pasaulio bankas sukūrė išpūdingą pasaulio rodiklių aibę, kurios apimtis per 30 lentelių, o ji skelbiama Pasaulio raidos ataskaitos (*World Development Report*) priede.

Ekonominių duomenų tikslumo tyrimai atskleidė, kad duomenų paklaida sudaro nuo 5 iki 20 procentų pateikiamų duomenų dydžio, atsižvelgiant į valstybę ir rodiklio kategoriją. Šią išvadą ekonomistas Oskar Morgstein 1963 m. apibūdino taip: „Ekonomika yra vieno skaitmens mokslas“.

Apibendrinus *socialinės statistikos objektą* galima apibrėžti kaip statistinių matavimų sistemos ir statistinių analizės metodų taikymą visuomenėje vykstantiems socialinės aplinkos reiškiniams ir procesams tirti bei jų kitimo laike tendencijoms įvertinti.

Statistikos uždaviniai ir funkcijos. Valstybėje statistika atlieka kelias svarbias funkcijas. Pirma, statistine informacija remiamasi formuojant vyriausybės politiką ir leidžiant įstatymus. Statistinis valstybės socialinės, ekonominės ir demografinės būklės įvertinimas parodo, ar reikia reguliuoti esamą padėtį ir priimti naujus ar pakeisti jau galiojančius įstatymus. Pavyzdžiui, nežinant vyrų ir moterų vidutinės gyvenimo trukmės, negali būti tinkamai nustatytas išėjimo į pensiją amžius, nežinant gyventojų pajamų pasiskirstymo, neįmanoma įvertinti skurstančiųjų sluoksnio, neįvertinus būsimo gyventojų skaičiaus ir sudėties, sunku planuoti mokyklų ir ikimokyklinių įstaigų tinklo pokyčius ir pan.

Antra, demokratinėje visuomenėje statistika kiekvienam piliečiui sudaro prielaidas susipažinti su valstybės socialine ir ekonomine raida bei galimybę vertinti valstybės institucijų veiklą. Svarbiausia sąlyga šiai funkcijai įgyvendinti yra kiekvieno teisė žinoti šalies socialinę, ekonominę, demografinę, ekologinę situaciją ir dalyvauti priimant sprendimus.

Trečia, statistinė informacija tradiciškai gali tarnauti kaip svarbus patriotinio išdidumo šaltinis. Istorikams žinoma XVII a. anglų ekonomisto William Petty knyga „Politinė aritmetika“ (*Political Arithmetic*), kurioje palyginami Anglijos, Prancūzijos ir Olandijos kariniai ir ekonominiai pajėgumai. Knygos tikslas buvo atskleisti akivaizdų Anglijos pranašumą šiose srityse.

Pastaraisiais dešimtmečiais pastebima vis didėjanti privačių organizacijų statistinės informacijos paklausa. Privačiame sektoriuje statistinė informacija naudojama teikiant kreditus, planuojant paklausą rinkoje ir pan. Kaip ir visi informacijos vartotojai, privatus sektorius turi skirtingų

tikslų ir skirtingų reikalavimų informacijos pateikimo greičiui. Pavyzdžiui, jei prieš valandą ar kelias valandas teisingai numatysime akcijų ar valiutos kurso kitimą, tai bus labai vertinga informacija. Tačiau ta pati informacija po kelių valandų jau bus bevertė. Informacijos vertės dydis priklauso ir nuo vartotojo tikslų.



SOCIALINIŲ ŽINIŲ STRUKTŪRA: DUOMENYS, MODELIAI IR TEORIJOS

„Ir štai žmonės eina: su nuostaba žiūri į aukštus kalnus, į tolimas jūras, į veržlias upes, į dangaus šviesulius, bet tuo pat metu užmiršta apie save pačius“.

Šv. Augustinas

XX amžiaus antroje pusėje socialiniuose moksluose vyravo loginio pozityvizmo idėjos ir požiūriai. Remiantis loginiu pozityvizmu, teorija – nagrinėjamos srities žinių apibendrinimas – turi būti pagrįsta svariais empiriniais apibendrinimais ar dėsniais. Išvados bei paaiškinimai yra atskirų faktų visuma, sudaryta remiantis bendresniu teiginiu, kuris savo ruožtu gaunamas remiantis dar bendresniu teiginiu ir t.t. Šiuolaikinė mokslo filosofija vis labiau prieštarauja klasikiniam loginio pozityvizmo požiūriui. Pagal šiuolaikinę mokslo filosofiją, mokslo dėsniai retai kada tiksliai atspindi realybę. Geriausiu atveju šie dėsniai yra idealizuotas sudėtingo ir sunkiai apibrėžiamo realaus pasaulio vaizdas. Iš esmės jie yra daugiau žmogaus proto kūrinys, o ne realaus pasaulio atspindys. Gierre [1999] pažymi, kad daugelis mokslinių dėsnių nėra universalūs ir net ne visada teisingi: „suprantami kaip visuotiniai tvirtinimai apie pasaulį, daugelis dėsnių, pretenduojančių į gamtos aiškinimą, iš esmės yra klaidingi... Tai yra dėl to, kad bet koks gamtos dėsnis apima tik keletą fizinių kiekybių, kai tuo tarpu gamta savyje turi daugybę kiekybių, kurios dažniausiai sąveikauja tarpusavyje, ir beveik nėra izoliuotų sistemų. Todėl negali būti daug realaus pasaulio sistemų, kurios tiksliai atitiktų suformuluotą gamtos dėsni“

Sprendžiamų ar analizuojamų socialinių problemų, susijusių su mokslinėmis žiniomis, jų gavimu ir vaizdavimu, aplinką sudaro keturi skirtingi, tačiau tarpusavyje susiję, elementai:

1. Realus pasaulis: šis pasaulis egzistuoja nepriklausydamas nuo žmonijos sukauptų žinių;

2. Teorija: aibė tarpusavyje logiškai ir nuosekliai susijusių sąvokų, apibūdinančių realaus pasaulio reiškinių, objektą ar jo dalį;
3. Duomenys: kurios nors realaus pasaulio sistemos stebėjimai ir matavimai;
4. Modelis: abstraktus, tačiau kruopščiai ir aiškiai apibrėžtas realybės reiškinių vaizdavimas, remiantis teorija arba duomenimis, arba ir teorija, ir duomenimis.

Žodis *teorija* nėra griežtai apibrėžtas, ir skirtingi žmonės dažnai jį supranta skirtingai, tiek kasdienėje kalboje, tiek ir mokslinėse diskusijose. Teorija yra realaus pasaulio supratimo išraiška, kuri gali būti pagrindas pasaulio reiškiniams ar objektams paaiškinti, prognozuoti ir reikšmingai kontroliuoti.

Duomenys sudaro ribotą realaus pasaulio dalies stebėjimų aibę. Reikia neužmiršti, kad beveik visuose sudėtingesniuose stebėjimų rezultatuose yra matavimų arba gavimo paklaidų.

Remiantis šiuolaikine mokslo filosofija, pagrindinis mokslo elementas yra *modelis*. Šiuo metu moksle dažniau naudojamas terminas „modeliu pagrįstas požiūris“ (*model-based view*), o ne senesnis filosofinis terminas „semantinis požiūris“ (*semantic view*). Modeliais pagrįstas požiūris nenubrėžia aiškios ribos tarp modelio ir teorijos. Teorija gali būti suprantama kaip teorinių modelių rinkinys. Taikant šias sąvokas, mokslo žinių prototipas yra ne empiriniai ar teoriniai dėsniai, bet modeliai kartu su išvardytomis realaus pasaulio sistemomis, kurioms šie modeliai gali būti taikomi. Modeliai iš esmės yra abstrakčios konstrukcijos, kuriomis mėginama pavaizduoti tik kai kurias realaus pasaulio savybes. Modeliais pagrįstas požiūris kreipia į abstrakčių modelių taikymą tiriant unikalius atvejus, kai tuo tarpu loginis pozityvizmas reikalauja atskirų atvejų empirinio apibendrinimo. Klausimas, ar modeliai yra empiriškai teisingi, neturi prasmės, kadangi to nė negali būti. Reikėtų klausyti ne apie modelio empirinį teisingumą, bet apie nagrinėjamo modelio tinkamumą konkrečiu empiriniu atveju. Su tuo yra susijęs ir klausimas, ar modeliai tam tikrais atžvilgiais atitinka nagrinėjamą realaus pasaulio dalį ir:

- ar jie yra pakankamai tikslūs;
- ar tinka tam tikru, tiksliai apibrėžtu, tikslu;
- ar aiškinamas reiškinys logiškai išplaukia iš modelio, pritaikyto konkrečiai realaus pasaulio daliai.

Giere pateikia pavyzdį – Žemės bei Mėnulio sistemos modelį, kuris

adekvačiai vaizduoja ir apskaičiuoja Mėnulio orbitą ir tinka norint paleisti raketą į Mėnulį, tačiau netinka norint aprašyti Veneros bei Žemės sistemą.

Akivaizdu, kad vienų modelių taikymo sritys yra platesnės nei kitų. Tais atvejais, kai modelių savybės vienodos, moksle pirmenybė teikiama tiems modeliams, kurie plačiau taikomi.

Išskiriami trys pagrindiniai modelių tipai:

1. Fiziniai modeliai (atomų modeliai, automobilio aerodinaminių savybių modeliai);
2. Vaizdiniai modeliai (įvairūs geologiniai žemėlapiai, demografinių pokyčių diagramos);
3. Teoriniai ir empiriniai modeliai (evoliucijos teorija ar Niutono gravitacijos teorija).

Fiziniai modeliai mažai tinka socialiniams mokslams.

Vaizdiniai modeliai yra perspektyvūs, jie vis plačiau taikomi, tačiau pateikiant sudėtingesnių procesų ar sistemų vaizdus, jų galimybės yra ribotos.

Teorinis modelis pateikia teorinių idėjų rinkinį. Teorinis modelis gali būti reiškiamas įprastine kalba, formaliomis loginėmis sistemomis, matematiškai, arba kompiuteriniais kodais. Toks modelis yra abstrakti konstrukcija, kuri konkrečiu tikslu gali būti tiek naudinga, tiek ir nenaudinga. Moksle neretai konkretus tyrimo tikslas yra reiškinių aiškinimas ar prognozė, o ne praktinis naudojimas. Šiuo atveju geriau tinka mažiau abstraktūs *empiriniai modeliai*. Empiriniai modeliai atspindi duomenų aibėje egzistuojančius ryšius ar struktūras. Šie modeliai mėgstami matematinio modeliavimo srityse. Tarp teorinių ir empirinių modelių nėra griežtos ribos, naudojama daug hibridinių arba pereinamojo tipo modelių. Empirinių modelių galima apibrėžti kaip „kilusį iš duomenų ir paremtą tikrai duomenimis“ (Edwards, Hamson, 1989). Ryšiai tarp kintamųjų čia nustatomi vertinant galimus kintamųjų matmenis ir jiems parenkant matematinę formą. Skiriamasis empirinio modelio bruožas yra tas, kad empiriniu modeliu vertinant kintamųjų ryšius neatsižvelgiama į teorines prielaidas ir nesiremiama teoriniais dėsniais. Dažnai empiriniai modeliai yra naudojami kaip daug sudėtingesnio modelio dalis. Kai nėra žinomų teorinių prielaidų, empirinis modelis padeda nustatyti ryšius tarp kintamųjų ir prognozuoti galimus pokyčius.

Naudojant modeliavimą, kuriame svarbiausią vietą užima individuali duomenų seka, teorijos indėlis yra mažas. Kraštutinis atvejis yra *ap-*

roksimacinės funkcijos, kurios taikomos ieškant funkcijos formos, geriausiai vaizduojančios dvišakes, trimates ar n -mates duomenų aibes, neatsižvelgiant į funkcijos ar jos parametru prasmę. Tipiniai aproksimacinių funkcijų naudojimo atvejai yra glodinimas, interpoliacija ar ekstrapoliacija. Skirtingai nei aproksimacinių funkcijų, *parametrinių funkcijų* parametrai gali turėti prasmines reikšmes. Parametrinių funkcijų naudojimas gali remtis teorija, o funkcijos forma ir parametrai atspindėti susijusią teoriją. Nesant teoriškai pagrįsto modelio, parametrai gali paprasčiausiai apibūdinti duomenų aibę. Šiuo atveju sistema, teikianti duomenis (priklusomą kintamąjį Y , susijusį su nepriklausomu kintamuoju X) yra traktuojama kaip juoda dėžė.

Daugiamatis statistinis modeliavimas panašus į teorinį parametrinį modeliavimą. Parametrai turi savarankiškas reikšmes, tačiau dažniausiai procesas, kurio metu iš įėjimo duomenų (X) gaunami išėjimo duomenys (Y) tiksliai neapibrėžiamas. Konkrečių kintamųjų parinkimas gali remtis teoriniais svarstymais, bet dažniau būna grindžiamas duomenų tinkamumu ar anksčiau nagrinėtais koreliaciniais ryšiais. Statistinis modelis vaizduoja duomenų struktūrą, kuri daugeliu atveju būna tiesinė. Tačiau statistinis modelis nėra grynai empirinis, visiškai apibrėžtas duomenimis. Visada turi būti išankstinės (*a priori*) prielaidos dėl kintamųjų įtraukimo į modelį ir funkcijos formos parinkimo. Toks modelis gali tikt geriausiai, tačiau pasirinkimas iš galimų modelių aibės yra labai ribotas. Daugiamaciai statistiniai modeliai nėra, ar bent nepretenduoja būti, realaus pasaulio sistemos, iš kurios duomenys buvo gauti, atvaizdas.

Kaip klasikinį teoriją pagrįsto modelio pavyzdį galima teikti Hernes (1972) pirmųjų vedybų tyrimą. Tyrimo pradžioje formuluojamos teorinės prielaidos: tam tikro amžiaus žmonių galimybės, arba teisingiau, tinkamumas vedyboms nėra vienodas, su amžiumi jis mažėja. Spaudimas tuoktis stiprėja didėjant susituokusiųjų procentui, tačiau tik iki tam tikro lygio, kol pradeda ryškėti tinkamų partnerių stygius. Šios teorinės idėjos buvo išreikštos asimetrinėmis logistinėmis matematinėmis funkcijomis, kurių eksponentinė tinkamumo tuoktis asimetrija su amžiumi mažėja. Parametrai įvertinami iš pirmą kartą susituokusiųjų dalies proporcijų pagal amžių. Parametrai vertinami pagal duomenis, o jų reikšmės prasmę įgyja pagal teorinius samprotavimus. Ši analizė gali būti pateikta kaip kontrastas aklam, mechaniniam matematinėms aproksimacinių funkcijų taikymui pirmųjų tuoktusių amžiaus duomenims arba vien tik teoriniams modeliams.

Šiame tyrime panaudota ganėtinai paprasta teorija ir nesudėtinga duomenų aibė. Hernes nebando sudaryti sudėtingą vedybų sistemos modelį, kuris turėtų įvertinti endogamiją (t. y. paprotį tuoktis su tos pačios visuomeninės grupės asmenimis), egzogamiją (santuokų tarp tos pačios giminės vyrų ir moterų draudimą), sutuoktinio parinkimą, ikisantuokinį pastojimą, sužieduotuves, povestuvinio gyvenimo aplinką ir kitus faktorius, turinčius įtakos pirmųjų vedybų amžiui.

Taigi, galima išskirti **tris pagrindines modelių rūšis**:

- modeliai, sudaryti remiantis ribotu teorinių idėjų rinkiniu. Visiškai įmanoma, kad gera teorija ar teorinis modelis gali geriau atitikti tikrovę nei duomenys ar statistiniai modeliai. Be to, į teorinius modelius gali būti įtraukti kintamieji, kuriuos žinome esant reikšmingus, tačiau kurių visai ar lengvai negalima išmatuoti. Teoriniuose modeliuose lengviau ir paprasčiau nagrinėti procesus su grįžtamaisiais ryšiais. Šie modeliai gali pateikti tikėtiną realaus pasaulio sistemų darbo vaizdą, kurį vargu ar gali pateikti statistiniai modeliai. Tačiau norint iširti sukurtos sistemos savybes ir parodyti, kad ši sistema tiksliai atskleidžia realų pasaulį, būtina turėti empirinius duomenis. Tik remiantis empiriniais duomenimis galima nustatyti koku tikslumu sukurta sistema atitinka nagrinėjamą realaus pasaulio dalį;
- modeliai, sudaryti iš ribotos duomenų aibės. Duomenys ir empiriniai modeliai yra arčiausiai realaus pasaulio. Šie modeliai leidžia nustatyti kintamųjų ryšius ir gali sudaryti gerą teorinių idėjų pagrindą.
- sudėtingi modeliai, aprėpiantys plačias teorines ir empirines žinias. Jie atspindi palyginti sudėtingą realaus pasaulio dalį ir pateikia ganėtinai tikslią jo kopiją. Šie sudėtingi modeliai žinomi kaip plataus masto modeliavimo (*large-scale simulation*). Paprastesniais atvejais, atsižvelgiant į tyrimo tikslus, šio tipo modeliuose gali būti apsieita ir su riboto dydžio teorinių idėjų ir duomenų aibėmis.

Visos šios modelių rūšys gali būti naudojamos ne tik realaus pasaulio sistemų aibėje egzistuojantiems ryšiams ar struktūroms paaiškinti, bet ir platesnio teorinio apibendrinimo pagrindui sukurti. Matematiniumi modeliavimu sudarytas abstraktus realios sistemos modelis leidžia suprasti realaus pasaulio sistemų veiklą, jas aiškinti ir prognozuoti artėjančius pokyčius.

1. SOCIALINIAI TYRIMAI IR VIEŠOJI NUOMONĖ

„Tas, apie kurį anksčiau kalbėjo, kad tai aš, dabar – jau kitas. O dabartinis aš, apie kurį dar niekas nežino, neišku kuo taps ateityje“.

Šao Jun

Viešąją nuomonę galima apibrėžti kaip „požiūrių ir lūkesčių rinkinį, tikėjimų konfigūracijas, kurios iškyla iš bendro viešųjų klausimų svarstymo ir apmąstymo“ (Goffman, 1969). Aiškiausiai viešoji nuomonė pasireiškia rinkimuose ir referendumuose. Lietuvos Respublikos Konstitucijos 4 straipsnis teigia: „Aukščiausią suverenią galią Tauta vykdo tiesiogiai ar per demokratiškai išrinktus savo atstovus“, o 9 straipsnio nuostata: „Svarbiausi Valstybės bei Tautos gyvenimo klausimai sprendžiami referendumu“. Konstitucijoje Tauta apibrėžiama kaip turinti valstybėje aukščiausią galią (suvereną), o išrinkti tautos atstovai privalo vykdyti tautos valią. Tautos valia yra abstrakti sąvoka, kuri praktiniame gyvenime dažniausiai keičiama sąvoka „viešoji nuomonė“.

Demokratinėje visuomenėje viešoji nuomonė turi lemiamą reikšmę formuojant valstybės politiką, sprendžiant socialines bei ekonomines valstybės problemas. Vertinant ir formuojant viešąją nuomonę būtina įvertinti ne tik visos valstybės, bet ir atskirų visuomenės sluoksnių bei socialinių grupių interesus. Visos visuomenės interesai gali nesutapti su atskirų socialinių grupių ar net sluoksnių nuomone. Atskirų asmenybių, kolektyvų ar gyventojų grupių nuomonę sąlygoja skirtingas kultūros lygis, išsilavinimas, turima informacija, todėl atskiros nuomonės ar socialinių bei ekonominių reiškinių vertinimas gali skirtis nuo daugumos socialinių grupių nuomonės. Visuomenės nuomonė formuojama pokalbių, diskusijų, ginčų metu.

Viešosios nuomonės tyrimai sudaro dalį visų socialinių tyrimų. Šių tyrimų metu kruopščiai tiriami žmonių požiūriai ir nuomonės socialiniais, politiniais ir kitais klausimais. Dažniausiai viešosios nuomonės tyrimai sukelia didelį visuomenės susidomėjimą, emocijas, būna plačiai

publikuojami ir aptariami. Kartais šių tyrimų rezultatai pateikiami iššaukiamai ar net tendencingai.

Viešosios nuomonės tyrimai dažnai atliekami prieš rinkimus, tiriant rinkėjų ketinimus, svarstant visuomenei aktualias problemas, pavyzdžiui, euro įvedimą, planuojant keisti viešąją tvarką ir pan.

Viešosios nuomonės tyrimus gali užsakyti politinės partijos, visuomeninės organizacijos, savivaldybės, vyriausybė, kitos organizacijos ir grupės, sprendžiančios socialinius, komercinius, politinius bei ideologinius klausimus.

Jeigu viešosios nuomonės tyrimo rezultatai yra publikuojami, tyrėjas privalo užtikrinti, kad spaudoje pateikti rezultatai kuo tiksliau atspindėtų viešąją nuomonę.

Viešosios nuomonės arba socialiniuose statistiniuose tyrimuose naudojami įvairūs apibrėžimai.

1.1. Apibrėžimai

Populiacija. Populiacija – tai aibė mus dominančių objektų (asmenų, daiktų ar sąrašų), iš kurių atrenkama imtis tyrimui. Pavyzdžiui, populiaciją gali sudaryti rinkėjai, studentai, gydytojai ir t.t.

Imtis. Imtis – tai populiacijos dalis, atrinkta statistiniam tyrimui, kurio tikslas gauti informacijos apie visą populiaciją.

Pagrindinės priežastys, dėl kurių dažniausiai atrenkama imtis, o ne tiriama visa populiacija:

- sutaupoma pinigų – imties tyrimas reikalauja mažiau sąnaudų. Praktiškai organizuojant tyrimus išlaidas rekomenduojama skirstyti į dvi dalis: viena pusė – duomenų rinkimui, kita pusė – statistinei analizei;
- sutaupoma laiko, pavyzdžiui, kandidatų į prezidentus reitingai po rinkimų nieko nedoms;
- kai kurie populiacijos nariai neprieinami, pavyzdžiui, atliekantys bausmę pataisos darbų įstaigose;
- kartais, norint ištirti atskirus elementus, tenka juos sunaikinti, pavyzdžiui, atliekant kokybės kontrolę, gaminį dažnai tenka sunaikinti;
- tikslumas. Informacija, gauta iš nekrupščiai ištirtos populiacijos, gali būti mažiau patikima nei gauta iš kruopščiai atrinktos imties.

Atranka. Atranka – tai metodas, kuriuo atrenkama tinkama imtis, turint tikslą nustatyti parametrus arba bruožus, būdingus visai populiacijai.

Prieš užsakant arba atliekant viešosios nuomonės arba socialinį statistinį tyrimą reikėtų detalčiai aptarti šiuos klausimus:

1. **Tikslai arba hipotezės.** Kokie yra tyrimo tikslai ir į kokius klausimus norima gauti atsakymą? Kokia populiacija bus tiriama, norint pasiekti tyrimo tikslų?
2. **Rezultatų pateikimo būdas.** Ar tyrimo rezultatai yra įdomūs visuomenei, ar jie bus publikuojami? Jei publikuojami, tai ar visi, ar tik atskiros tyrimo dalys?
3. **Statistinis modelis.** Ar planuojamas statistinis eksperimentas (t.y. bus renkami pirminiai duomenys), ar bus naudojami jau atliktų tyrimų duomenys (t.y. antriniai duomenys)? Kaip bus atrenkama imtis? Ar galimi tokie atrankos būdai, kuriais atrinkta imtis būtų netipinė ir nereprezentatyvi? Jei taip, tai kokių atsargumo priemonių bus imtasi? Ar bus sudaroma kontrolinė grupė? Jei bus atliekama apklausa, ar apskaičiuoti apklausos kaštai ir apklausos trukmė?
4. **Tyrimas.** Ar aiškūs kintamųjų (požymių) apibrėžimai, jų klasifikavimas, matavimo skalės? Ar klasifikavimo ir matavimo metodai suderinti su tyrimo tikslais? Ar tyrimas patikimas ir galės būti pakartotas (jei reiktų apginti jūsų išvadas)?
5. **Analizė.** Ar yra pakankamai duomenų pagrįstai statistinei analizei atlikti? Jei taip, ar statistinės analizės metodų sąlygos tinka duomenų šaltiniui ir atitinka duomenų pobūdį? Analizė turi būti tinkamai atlikta ir paaiškinta.
6. **Išvados.** Kurios išvados atitinka pradines hipotezes? Kurios neatitinka? Ar išvados dera tarpusavyje?
7. **Rezultatų pateikimas.** Gauti rezultatai turi būti pateikti aiškiai, objektyviai, kvalifikuotais, bet ne techniniais terminais ir detalėmis. Ar analizės dalis galima pakartoti?
8. **Vadovų (vertintojų) išvados.** Jei gauti rezultatai ir teikiamos rekomendacijos nėra aiškiai išdėstytos, arba jei netinkamai aprašytos, jos būna nesuprantamos vertintojams ir sprendimų priėmėjams. Tokiu atveju rezultatai atrodo nepagrįsti ir neįtikinami.

Pirmame statistinio tyrimo etape gaunami pradiniai statistiniai duomenys sudaro būsimo statistinio „pastato“ pagrindą. Jeigu renkant

pradinius duomenis padaroma klaida arba surenkama nekokybiška medžiaga, tai turi didelės įtakos būsimų išvadų teisingumui ir pagrįstumui. Ne kiekvienas duomenų rinkinys yra statistinis tyrimas. Apie statistinį tyrimą galima kalbėti tik tada, kai tiriami statistiniai dėsningumai, t.y. tokie, kurie atsiskleidžia tik tiriant daug nagrinėjamos aibės elementų.

Pradiniai duomenys yra tik neapdorota informacija ir patys savaime jie nėra žinios. Seka, kuria iš duomenų gaunamos žinios, yra ši: *duomenys* → *informacija* → *faktai* → *žinios*. Duomenys tampa *informacija*, kai jie tiesiogiai susiejami su sprendžiama problema. Informacija tampa *faktu* tada, kai duomenys ją patvirtina. Faktai yra tai, ką atskleidžia duomenys. Faktai tampa *žiniomis*, kai jie naudojami sprendimo procesui sėkmingai užbaigti. Žinios, kurios remiasi patyrimu ir išmintimi, sudaro galimybę priimti geriausią sprendimą, optimaliai panaudoti laiką ir turimus resursus. Didėjant sudaryto statistinio modelio tikslumui gerėja ir sprendimo priėmimo procesas.

1.2. Statistinės informacijos rinkimo būdai

Visą statistinę informaciją sąlyginai galima suskirstyti į dvi pagrindines grupes:

- oficiali apskaita, kurią organizuoja ir vykdo oficialios, dažniausiai valstybinės, institucijos;
- specialiai organizuojami statistiniai tyrimai.

Statistiniuose tyrimuose kartais naudojami ir abu apskaitos metodai, kadangi papildo ir patikslina duomenis apie tyrimo objektą.

Specialiai organizuoti statistiniai tyrimai naudojami tada, kai reikia informacijos, kurios oficiali statistika neteikia. Pagrindiniai tyrimo metodai yra šie:

- apklausa;
- socialinis stebėjimas;
- socialinis eksperimentas.

Plačiausiai taikomas apklausos metodas. Taikant šį metodą, per trumpą laiką ir su nedidelėmis lėšų sąnaudomis galima apklausti tiriamos populiacijos reprezentatyvios imties elementus. Apklausos metodas lengvai formalizuojamas – tai palengvina surinktų duomenų analizę.

Atliekantys apklausą tyrėjai skirstomi į dvi grupes:

1. **Atsakingieji tyrėjai**, kurie numato tyrimo metodologiją (sudaro apklausos klausimus ir pan.) ir sudaro bei kontroliuoja tyrimo planą. Šie tyrėjai gali turėti įtakos tyrimo rezultatams ir duomenų publikavimui.
2. **Tyrimo vykdytojai arba interviuotojai (apklausėjai)** – asmenys, kurie renka faktinę medžiagą. Šie tyrėjai vykdo tyrimo planą, renka ir apdoroja duomenis. Jie nesudarinėja klausimų ir neinterpretuoja rezultatų. Tačiau jie gali naudoti savo įtaką tam, kad užtikrintų deramą tyrimo vyksmą, laikantis vykdymo plano ir teisės aktų, reglamentuojančių duomenų apsaugą.

Apklausiai atlikti naudojamos trys pagrindinės formos:

- pašnekesys;
- interviu;
- anketavimas.

Pašnekesys su asmenimis, kurių nuomonė tiriama, vyksta laisvai bendraujant pagal kruopščiai apmąstyta planą. Pašnekesio sėkmė – išankstinis pasirengimas ir gebėjimas sukurti psichologinį kontaktą su respondentu. Dažnai pašnekesio metu tenka liesti gyvenimo sritis, kurių nagrinėjimas respondentui ne visada naudingas. Todėl kartais rekomenduojama pokalbio neįrašinėti, tik vėliau atkurti jį iš atminties. Apklausiant šiuo metodu, reikia nusiteikti ilgam, kruopščiam procesui, kurio rezultatai sunkiai formalizuojami. Šis metodas taikomas teismuose, advokatūroje, jį taiko gydytojai ir žurnalistai.

Interviu – taip pat yra pašnekesys, tik jo metu laikomasi tiksliai apibrėžtų klausimų. Interviu trumpesnis, ir gvildenamas apibrėžtas klausimų skaičius. Interviu metu būna mažiau galimybių psichologiniam kontaktui užmegzti, ir tai gali turėti įtakos rezultatams. Interviu skirstomas į keletą pagrindinių tipų: laisvas, formalizuotas (standartizuotas), periodinis, grupinis.

Laisvas interviu panašus į pašnekesį, kai respondentui pateikiama klausimų, kurių pagrindu vyksta pašnekesys.

Formaliame interviu respondentui pateikiama griežtai suformuluotų klausimų ir stengiamasi, kad jų forma būtų kuo vienodesnė. Pageidautina atsakyti „taip“ arba „ne“, arba vertinama balais (1,2,3...). Tokiu būdu gauti interviu rezultatai yra palyginami ir lengvai analizuojami.

Periodiniame interviu su tais pačiais asmenimis (arba vienodo tipo grupėmis) bendraujama nustatytais laiko tarpais. Taip ištiriami nuomonių ir pažiūrų pokyčiai.

Grupiniame interviu sužinoma grupės nuomonė, kuri ne visada sutampa su daugumos nuomone, ir dažniausiai neatspindi atskirų individų nuomonės. Dažniausiai taip sužinoma grupės lyderių nuomonė, požiūris bei orientacija.

Anketavimas yra plačiausiai taikomas ir sudaro tarsi paskutinį apklausos laiptelį su dar griežtesnėmis apklausos taisyklėmis. Anketą sudaro keletas tarpusavyje susijusių klausimų, į kuriuos reikia gauti respondentų atsakymus. Pati anketa nėra griežtos formos. Sakoma, kad anketos sudarymas yra menas. Galima pateikti tik bendrus reikalavimus. Klausimų turinys, skaičius ir eilė priklauso nuo tyrimo tikslų. Prieš pradėdant sudarinėti klausimus reikia tiksliai formuluoti: a) ką jūs norite sužinoti ir b) apgalvoti, ar populiacija, kurią jūs planuojate apklausti, nori jums pateikti informacijos. Klausimuose rekomenduojama naudoti tik gerai visiems žinomus žodžius, nevartoti sudėtingų struktūrų.

Dažniausiai naudojamos kelios anketinės apklausos rūšys:

- anketinė apklausa el. paštu arba paštu;
- telefoninė apklausa;
- apklausa per masines informacijos priemones;
- ekspertų apklausa, kai atrenkama žmonių grupė (ekspertai), turinti kurios nors srities žinių. Ekspertai analizuoja problemą ir kiekybiškai ar kokybiškai ją įvertina.

Anketinės apklausos el. paštu arba paštu pagrindinis trūkumas – mažas atsakiusių skaičius (nuo 20 iki 40 proc.). Atsakymų skaičiui padidinti taikomi įvairūs metodai. Vienais iš jų dėmesys kreipiamas į anketos pateikimą – siunčiant būtina įdėti voką su ženklu užpildytai anketai atsiųsti. Be to, siunčiami vokai ir anketos turi būti geros kokybės (popierius, įdomūs pašto ženklai ir pan.). Taikant kitus metodus, rekomenduojama skatinti – paskatinimai dažniausiai siunčiami kartu su pirmuoju laišku. Paskatinimų gali būti įvairių: kanceliarinės prekės, knygos, kalendoriai ir pan. Atsakomų anketų skaičių didina ir primenamieji laiškeliai, kurie siunčiami praėjus 2–3 savaitėms nuo anketų išsiuntimo.

Telefoninė apklausa dėl besiplečiančių telefoninių ryšių naudojama vis dažniau. Šis būdas turi daug privalumų – atsakymai gaunami iškart, apklausa yra pigi ir atsakiusių būna pakankamai daug. Telefoninė apklausa dažniausiai trunka 10–15 min.

Apklausa per masines informacijos priemones – laikraščiuose ir žurnaluose spausdinamos anketos. Taip tiriama skaitytojų nuomonė. Šio

tipo apklausos, kaip ir apklausos paštu, problema – mažas atsakiusių skaičius.

Ekspertų apklausai sudaroma specialistų grupė. Ekspertų apklausos metodas taikomas, kai problemai ar situacijai įvertinti reikia specialių žinių, pavyzdžiui, įvairių institucijų veiklai įvertinti. Čia susiduriama su ekspertų parinkimo problema. Ekspertai turėtų būti objektyvūs ir beveik vienodos kompetencijos.

1.3. Klausimų formulavimas ir apklausos lapų sudarymas

Atsakingasis tyrėjas privalo užtikrinti, kad sudaryti klausimai atskleistų tyrimo tikslus ir neturėtų įtakos respondento nuomonei. Klausimai turi būti sudaryti taip, kad neskatintų respondento rinktis kurį nors atsakymą. Galima taip sudaryti sakinius ar sakinių grupes, kad būtų skatinama rinktis arba atmesti tam tikrą požiūrį ar nuomonę.

Sudarant klausimus atsakingasis tyrėjas privalo užtikrinti, kad:

- klausimai atitiktų tyrimo tikslus;
- klausimų formulavimas būtų tinkamas ir suprantamas žmonėms, kurie bus apklausiami;
- respondentai gebėtų atsakyti į klausimus taip, kad kuo geriau atspindėtų savo norimą pareikšti požiūrį;
- klausimų negalima būtų interpretuoti dviprasmiškai.

Sudarius klausimus dar kartą reiktų patikrinti ar:

- respondentai supras klausimus?
- noriai atsakinės į klausimus?
- gebės atsakyti į klausimus?

Viename klausime turėtų būti tik vienas objektas. Dažnai (bet nebūtinai!) anketose būna šios dalys: įvadinė, padėties, pagrindinė ir baigiajanti.

Įvadinėje anketos dalyje formuluojamas kreipimasis į respondentą, pateikiamos apklausos sąlygos ir atsakymų teikimo taisyklės.

Padėties dalį sudaro klausimai apie socialinę – demografinę respondento padėtį. Šie klausimai skirti analizės rezultatų interpretavimui palengvinti.

Pagrindinę dalį sudaro klausimai apie nagrinėjamą temą.

Baigiamojoje dalyje respondentui gali būti suteikta galimybė laisvai pasisakyti apie tiriamą problemą.

Padėties ir baigiamoji dalys nėra privalomos, jos gali būti praleistos. Įvadinė dalis gali būti pakeista žodiniu paaiškinimu.

Pagrindinės anketos dalies klausimai skirstomi į tris grupes:

- dichotominiai,
- atvirieji,
- uždarieji.

Dichotominiai klausimai yra tokie, į kuriuos galima atsakyti tik „taip“ arba „ne“ Kartais jie patikslinami papildomu klausimu: „Jei taip, tai kodėl?“ Į šį klausimą tikimasi savarankiško atsakymo.

Atvirieji klausimai numato laisvos formos atsakymą. Pateikiant šio tipo klausimus galima gauti visą informaciją, tačiau atsakymus sunku analizuoti ir formalizuoti. Todėl didelės apimties apklausoje tokių klausimų turėtų būti minimaliai.

Uždarieji klausimai numato daugelį parengtų atsakymų. Parenkamas tas, kuris atitinka respondento nuomonę. Rekomenduotina ne mažiau kaip keturi atsakymų variantai. Kartais anketose pateikiama uždarytųjų klausimų, į kuriuos atsakymus reikia sunumeruoti pagal svarbą. Pavyzdžiui: „Kokios savybės, esančios sąrašė, neleistinos valstybės tarnautojams?“ Toliau stulpelių abėcėlės tvarka nurodomos savybės: „, kategoriškumas, neryžtingumas, netaktiškumas, šiurkštumas“ ir t.t. Labai svarbu atsakymų skalė. Pavyzdžiui, pateikus klausimo: „Ar sėkmingai jums klostėsi gyvenimas?“ atsakymus skalėje nuo 0 iki 10, galima tikėtis iki 35 proc. teigiamų atsakymų. Pakeitus atsakymų skalę nuo –5 iki 5, teigiamų atsakymų gali sumažėti iki 20 procentų.

Privatūs duomenys. Tai duomenys, į kuriuos respondentai gali jautriai reaguoti ar nenorėti jų atskleisti. Tai šie duomenys:

- rasė ar etninė grupė;
- politinė nuomonė;
- religinis arba panašus tikėjimas;
- narystė profesinėje sąjungoje;
- fizinė arba psichinė sveikata;
- seksualinis gyvenimas;
- įrodyti arba įtariami nusikaltimai;
- nusikaltimai, perduoti nagrinėti teismui.

Tirdami viešąją nuomonę tyrėjai dažnai stengiasi gauti atsakymus ir į šiuos klausimus. Šių subtilių sričių klausimai turėtų būti pateikiami labai jautriai arba netgi netiesiogiai. Respondentams visada turi būti palikta teisė neatsakyti į šiuos klausimus ir/arba nutraukti apklausą šioje vietoje.

Netiesioginio klausimo formulavime dažniausiai vertinami kitų žmonių poelgiai ar nuomonės. Šie klausimai dažnai prasideda frazėmis: „Kaip jūs manote,..“, „Kaip jūs vertintumėte,..“ ir pan.

Klausimus galima skirstyti ir kitaip, pavyzdžiui:

- aprašomieji,
- siejamieji,
- priešastiniai.

Aprašomieji klausimai pateikiami tada, kai norima gauti duomenų apie esamą padėtį arba išsiaiškinti vykstančius procesus. Pavyzdžiui, apie kasdienį gyvenimą galima paklausti: „Ar esate patenkintas savo gyvenimo sąlygomis?“, „Ar jus tenkina jūsų pajamos?“, „Ar jūs jaučiate pasididžiavimą savo valstybės karinėmis pajėgomis?“, „Kaip jūs manote, ar gyvenimas artimiausiu metu pagerės?“ ir pan.

Siejamaisiais klausimais stengiamasi sužinoti respondento nuomonę apie galimą dviejų reiškinių ryšį, pavyzdžiui: „Kaip jūs manote, ar blogai sutariantis su savo žmona asmuo turi teisę siekti ištuokos?“, „Ar tiesa, kad berniukai lengviau išmoksta matematikos, o mergaitės – kalbų?“ ir pan.

Priežastiniais klausimais stengiamasi įvertinti galimą keleto veiksnių įtaką procesui ar konkrečiam rezultatui, pavyzdžiui: „Kas lėmė jūsų pasirinkimą dėl vienos ar kitos verslo valdymo programos: a) programos kaina; b) programos galimybės; c) programos gamintojų teikiamas aptarnavimas?“.

1.4. Tyrimo rezultatų analizė ir ataskaitos pateikimas

Iš statistiko ataskaitos. „Tiriant vaistų poveikį pelėms, buvo gauti šie rezultatai: trečdalis pelių teigiamai reagavo į vaistus ir pasveiko; trečdaliui pelių vaistai nepadarė, ir jos nugaišo, trečia pelė pabėgo...“

Socialinio tyrimo svarumas ir vertė priklauso nuo trijų pagrindinių aplinkybių:

- nuo tyrėjų arba atliekančios tyrimą organizacijos objektyvumo ir kvalifikacijos;
- nuo taikytų analizės metodų ir jų efektyvumo;
- nuo rezultatų pateikimo ir naudojimo būdo.

Jeigu viešosios nuomonės tyrimai yra publikuojami viešai, prie jų turi būti pridėta oficiali tyrimo suvestinė, o joje turi būti nurodyta:

1. Tyrimą atlikusios organizacijos pavadinimas;
2. Medžiagos rinkimo laikas;
3. Apibrėžta tiriamoji populiacija;
4. Naudotas atrankos metodas, imties dydis ir apklausos vieta. Turi būti nurodyta, kuris regionas buvo tiriamas, kaip jis buvo suskirstytas smulkiau ir kaip imtis išdėstyta geografiškai.
5. Apklausos būdas ir sutikusių dalyvauti tyrime respondentų procentas. Jei šis rodiklis yra mažas, reiktų išsiaiškinti nenoro dalyvauti tyrime priežastis.
6. Norint išvengti galimų nesusipratimų, turėtų būti paskelbtas teiktų klausimų sąrašas.
7. Jeigu pateikiami rezultatai labai skiriasi nuo pradinių surinktų duomenų, turi būti nurodyta, ar buvo taikyti statistiniai svorių metodai bei kitos statistinės procedūros.

Tinkamos formos ataskaitos publikavimo pavyzdžiai:

1. *Asahi Shimbun* užsakymu, tyrimą Japonijoje 1998 m. spalio 4 ir 5 dienomis atliko *Vox Populi*. Tyrimui buvo atrinkta 3000 balsavimo teisę turinčių gyventojų. Atsakymus pateikė 74 procentai arba 2229 asmenys. Tyrimui buvo sudaryta atsitiktinė sluoksniinė dviejų lygių imtis. Visa valstybė buvo padalinta į 348 sluoksnius pagal administracinį padalijimą. Iš kiekvieno sluoksnio buvo atrinkta viena rinkiminė apylinkė. Iš šios apylinkės buvo atrinkti vidutiniškai 9 respondentai.

2. Šis tyrimas buvo atliktas telefonu, *ABC Research Ltd* agentūros užsakymu. Tikimybinė reprezentatyvi imtis iš 1002 suaugusių asmenų, vyresnių nei 18 metų, buvo sudaryta atsitiktinės atrankos metodu. Atsakymus pateikė 68 procentai arba 681 asmenys. Tyrimas buvo atliekamas 2002 m. gegužės 1–4 dienomis Klaipėdoje.

Pastaba: Tais atvejais, kai tyrimo rezultatams laikraščiuose ir žurnaluose skiriamos kelios eilutės, dažnai tyrimo techninės detalės nenurodomos. Šiuo atveju būtina nurodyti, kur galima rasti visą informaciją.

1.5. Imčių sudarymo būdai

Norėtuši taip sudaryti imtį, kad būtų minimali ir paklaida, ir sąnaudos. Deja, paprastai paklaidų mažinimas didina sąnaudas. Tačiau minimalią paklaidą galima gauti tinkamai atrinkus elementus. Imtys sudaromos įvairiais būdais, tačiau pagrindiniai imčių sudarymo būdai yra du: *tikimybinis* ir *netikimybinis*. Atrenkant tikimybinio būdu, kiekvieno elemento tikimybė patekti į imtį yra žinoma, o sudarant netikimybinės imtis, daug lemia atsitiktinumas.

Netikimybinės imtys. Iš netikimybinio imčių sudarymo būdų populiariausi yra šie: *ekspertinė imtis*, kai elementai į imtį įtraukiami atsižvelgiant į ekspertų nuomonę; *kvotinė imtis*, kai iš anksto nustatomos atskirų populiacijos dalių kvotos; *proginė imtis*, kai į imtį įtraukiami patogiausiai ištiriami elementai. Taip sudarytos imtys dažniausiai būna ne-reprezentatyvios.

Tikimybinės imtys. Populiariausia yra *paprastoji atsitiktinė imtis* (*atsitiktinės atrankos* metodas). Atrenkant elementus šiuo metodu, visų populiacijos elementų galimybės patekti į imtį turėtų būti vienodos. Šis metodas realizuojamas dviem būdais. Pirmuoju būdu elementai numeruojami, maišomi, dedami į dėžę ir iš jos traukiami nesirenkant. Antruoju būdu kompiuteriu generuojami atsitiktiniai skaičiai, kurie yra atrenkamų elementų numeriai.

Praeityje, dažnai ir dabar, imtis sudaroma mechaniškai, atrankos nesiejant su atsitiktinių skaičių gavimo procedūromis. Taip sudaryta imtis vadinama *sistemingąja imtimi*. Šiuo būdu vienodu žingsniu atrenkamas kas kelintas populiacijos elementas. Pavyzdžiui: jei populiaciją sudaro 100 000 elementų, o reikia atrinkti 1000 elementų, tai į ją patenka kiekvienas šimtas elementas. Pirmąjį elementą stengiamasi parinkti atsitiktiniu būdu arba jis imamas iš pirmosios imties vidurio. Jei populiacija ganėtinai didelė, šis atrankos būdas artimas atsitiktinei atrankai.

Sluoksninė imtis naudojama tuo atveju, kai populiacija nėra vienalytė, o gali būti padalinta į mažesnes aibes (sluoksnius), kurių kiekviena yra vienalytė tiriamojo požymio atžvilgiu. Jei iš populiacijos, kurios dy-

dis N , norima atrinkti imtį iš n elementų, o populiaciją sudaro k vienalyčių sluoksnių, tai i -tojo sluoksnio dydis žymimas N_i , ir sluoksninių imčių dydžiai apskaičiuojami taip: $n_i = n(N_i/N)$, čia: $i=1,2,\dots,k$. Apibendrinant gautą informaciją, reikia atsižvelgti į sluoksnyje esamos populiacijos dalies dydį. Dažnai padalinti populiaciją į sluoksnius, juolab nustatyti sluoksnio dydį, nėra paprasta.

Lizdinė imtis naudojama tada, kai populiacija yra vienalytė, tačiau gali būti padalinta į mažesnes aibes. Daugelyje taikymų ši dalijimą lemia atstumas. Pavyzdžiui: tiriant savivaldybes, ligonines, koledžus, kurių yra visoje šalyje. Šiuo atveju, atrinkus elementus atsitiktinės atrankos būdu, kelionėms tektų skirti papildomai išlaidų, todėl atsitiktine atranka dažniausiai parenkamos atskiros aibės. Į imtį patenka visi atrinktųjų aibių elementai.

Praktikoje imtys sudaromos ir kitais būdais – įvairiai derinant minėtus imčių sudarymo būdus.

Neatsakymo lygis. Apklauskos paklaidoms labai didelės įtakos turi atsakymų stygius (daug neužpildytų anketų, respondentai atsisako dalyvauti tyrime). Pavyzdžiui: 2001 m. Lietuvos namų ūkio biudžeto tyrimui buvo atrinkta 10 tūkstančių namų ūkių, iš kurių 24 procentai atsisakė dalyvauti tyrime. Penkiuose didžiausiuose Lietuvos miestuose atsisakė dalyvauti 36 proc., kaimuose – 13 proc. Neatsakymo lygiui nustatyti naudojamosi vadinamuoju atsakymo lygiu: atsakymo lygis – tai atsakiusių respondentų skaičius, padalintas iš visų parinktų respondentų skaičiaus.

Imties paklaidos. Visada tikimasi, kad imtis atskleis populiaciją, iš kurios ji atrinkta, tačiau nėra jokių garantijų, kad imtis tiksliai reprezentuoja populiaciją. Dažniausiai imtį nereprezentatyvią padaro atrankos klaidos. Pagrindinės atrankos klaidų priežastys yra dvi:

- *Atsitiktinumai.* Pavyzdžiui: išmatavus šimto atrinktų moterų ūgį, visos buvo ne mažesnės nei 170 cm. Tai nelabai tikėtina, bet reisykiais taip gali atsitikti.
- *Tendencingumas.* Tendencingai atrenkami elementai su tam tikromis savybėmis. Ši klaida paprastai yra neteisingo atrankos plano rezultatas. Pavyzdžiui: apklausiant per internetą, bus sužinota nuomonė tik tos dalies gyventojų, kurie naudojami interneto ryšiu.

Atsakymų matavimo klaidos. Šių klaidų priežastis yra būdas, kuriu matuojama. Yra žinoma istorija, kai prancūzų astronomas pasiūlė naują teoriją, kuri rėmėsi atskiros žvaigždės skleidžiamo šviesos spektro tyrimais. Astronomas rėmėsi savo originaliais stebėjimais. Patikrinus

paaikškėjo, kad jo matavimo instrumentas buvo užterštas cigarečių dūmais. Dažniausiai pasitaikančios matavimo klaidos:

- *Apklaustos efektas*. Tas pats asmuo skirtingų apklausų metu gali pateikti skirtingus atsakymus. Pavyzdžiui: paklausus kiek respondentui metų, kai kurie asmenys, norėdami atrodyti jaunesni, linkę „numesti“ keletą metų, o kai kurie – vyresni, linkę nurodyti daugiau metų nei yra iš tikrųjų. Tikslesni rezultatai gaunami klausiant gimimo metų.
- *Atsakančiojo (respondento) efektas*. Kai kurie respondentai pateikia neteisingus atsakymus, norėdami padaryti įspūdį klausiančiajam. Pavyzdžiui: žinoma atveju, kai apklausiant ūkininkų poras apie gautą derlių, vyras ir žmona pateikia labai skirtingus duomenis.
- *Nepasitikėjimas tyrimo tikslu*. Tyrimo tikslo žinojimas kartais gali išprovokuoti neteisingus atsakymus. Klasikinis pavyzdys gali būti klausimas apie pajamas. Jei šį klausimą užduoda oficialus asmuo, atsakymas gali ženkliai skirtis nuo tikrųjų pajamų.
- *Tendencinga apklausa*. Jei tyrimą atlieka suinteresuoti asmenys, apklausa arba tyrimas gali būti atliekami tendencingai. Pavyzdžiui: tiriant trąšų poveikį derliui, gali būti parenkami geresni sklypai.

1.6. Matavimų skalės

Socialinių statistinių tyrimų kintamieji skirstomi į *kokybinius* ir *kiekybinius*. Savo ruožtu, kokybiniai kintamieji gali būti matuojami pagal *pavadinimų* (arba *nominaliąją*) bei *ranginę* skalę. Kiekybiniai kintamieji matuojami pagal *intervalinę* bei *santykinę* skalę.

Tipiškas *pavadinimų* skalės kintamasis – lytis. Galima skaičiais 1 ir 2 paženklinėti atitinkamai vyrą ir moterį, arba ženklus sukeisti vietomis, arba paženklinėti kitais skaičiais (kodais). Šie kodai tarpusavyje nelyginami. Jei kintamasis išmatuotas pagal pavadinimų skalę, jo statistinei analizei taikomų metodų rinkinys labai ribotas. Pagrindinė tokių kintamųjų analizė – dažnių analizė. Pavadinimų skalės kintamiesiems dažnai taikomi grupavimo metodai, kuriais turima imtis padalinama į atskiras kategorijas.

Dažnai kintamuosius galima vertinti pagal iš anksto nustatytus prioritetus – rangus. Pavyzdžiui: kalbant apie aptarnavimo kokybę, ją galima suskirstyti: puiki, gera, vidutinė, prasta. Koduojant šiuos duomenis, svarbu sąraše nurodyta eilė. Šiuos duomenis galima palyginti tarpusavyje: puiki kokybė aukštesnė nei gera, gera kokybė aukštesnė nei vidutinė ir t.t. Šie kintamieji yra *ranginės* skalės. Nors kintamųjų reikšmės palyginamos tarpusavyje, negalima teigti kad skirtumas tarp puikios kokybės ir geros yra toks pat, kaip tarp vidutinės ir prastos. Prie klasikinių ranginės skalės kintamųjų priskiriami kintamieji, gauti siejant duomenis į klases, pavyzdžiui: pajamas skirstant intervalais: iki 500 Lt; nuo 501 iki 1000 Lt; nuo 1000 iki 2000 Lt; per 2000 Lt.

Ranginės skalės kintamiesiems be dažnių analizės leidžiama apskaičiuoti kai kurias statistines charakteristikas, pavyzdžiui, medianą. Kai kuriais atvejais galima apskaičiuoti vidurkį. Jeigu reikia nustatyti ryšį (koreliacija) su tokio pat tipo kitu kintamuoju, taikomi ranginės koreliacijos koeficientai. Norint palyginti skirtingas ranginėje skalėje išmatuotų kintamųjų imtis, taikomi neparamestriniai kriterijai, kuriuos skaičiuojant naudojami rangai.

Jei ranginio kintamojo skalę sudaro 15 ir daugiau pozicijų, tai kintamąjį galima nagrinėti kaip kiekybinį.

Kiekybiniai (skaitiniai) kintamieji gali būti matuojami pagal intervalinę arba santykinę skalę. Prie *intervalinių* kintamųjų priskiriama temperatūra, intelektinis koeficientas, kalendorinis laikas. Jie ypatingi tuo, kad nulinis taškas yra laisvai parenkamas. Santykinė skalė skiriasi nuo intervalinės tik tuo, kad joje yra fiksuotas nulinis taškas. Atliekant statistinę analizę, skirtumai tarp intervalinės ir santykinės skalės kintamųjų paprastai laikomi nereikšmingais, ir šių kintamųjų analizei taikomi tie patys metodai. Dauguma metodų, taikomų intervalinių ir santykinų kintamųjų analizei, remiasi prielaida, kad šių kintamųjų skirstinys yra normalusis. Esant tokiam skirstiniui, dauguma reikšmių telkiasi apie vidurkį, o tolstant nuo jo į abi puses, stebėjimų tankis tolygiai simetriškai mažėja.

1.7. Imties dydis

Prieš nutariant, kokio dydžio imtis bus naudojama tyrimui, reikia apibrėžti populiaciją, pavyzdžiui: visi Kėdainių miesto vyresniųjų klasių moksleiviai. Šiuo atveju teks sudaryti visų Kėdainių miesto vyresniųjų klasių moksleivių sąrašą, ir tik tada numatyti imties dydį.

Imties dydžio nustatymo problema yra viena iš sunkesniųjų. Imties dydį gali sąlygoti įvairūs apribojimai, pavyzdžiui: turimos tyrimui lėšos gali turėti didelės įtakos imties dydžiui. Kai tyrimui lėšos yra patvirtintos, planuojant išlaidas naudinga vadovautis paprasta taisykle: pusę lėšų išleisti duomenų rinkimui, o kitą pusę – duomenų analizei. Šis apribojimas turi įtakos tiek imties dydžiui, tiek tyrimo planavimui ir duomenų rinkimui.

Apskritai, imties dydis priklauso nuo būsimos analizės tipo; nuo norimo pasiekti įvertinimo tikslumo; nuo vienu metu tiriamų kintamųjų skaičiaus; nuo tiriamos populiacijos vienalytiškumo. Pavyzdžiui: jei statistinio eksperimento pagrindinis tikslas yra eksperimentiškai įvertinti vidurkius ir palyginti jų skirtumus, gali pakakti šimto elementų imties.

Svarstant matematiškiau, imties dydis yra funkcija, priklausanti nuo kelių argumentų: norimo pasiekti įverčių tikslumo, populiacijos sklaidos ir norimo patikimumo.

1 pavyzdys. Tarkime, norima įvertinti populiacijos narių vidutinį uždarbį. Prieš atsakant į šį klausimą, pirmiausia reikia apibrėžti šią informaciją:

- 1) koks bus pasiklovimo lygmuo,
- 2) kokių tikslumu norima įvertinti parametą.

Pavyzdžiui: norima įvertinti atlyginimų vidurkį ± 10 Lt tikslumu per mėnesį, pasiklovimo lygmuo – 95 procentų. Imties dydį n galima nustatyti pagal pasiklovimo intervalo formulę:

$$\bar{x} \pm 1,96 \frac{s}{\sqrt{n}} .$$

Čia: s – standartinis nuokrypis.

Standartizuoto normaliojo skirstinio 95 proc. pasiklovimo lygmenį formulėje atitinka kvantilio reikšmė 1,96. Reikšmes, atitinkančias pasirinktą pasiklovimo lygmenį, galima rasti statistinėse lentelėse. Kadangi

pasirinktas tikslumas yra ± 10 , arba intervalo plotis 20 Lt, tai dydis $1,96 \cdot s / \sqrt{n}$ turi būti 10. Nors aritmetinio vidurkio \bar{x} reikšmė yra nežinoma, tačiau matyti, kad šis dydis įtakos sprendimui neturi, kadangi norima gauti intervalą, nepriklausomai nuo \bar{x} reikšmės. Taigi reikia išspręsti lygtį:

$$10 = 1,96 \frac{s}{\sqrt{n}} \text{ arba kitaip } \sqrt{n} = 1,96 \frac{s}{10}.$$

Nežinoma standartinio nuokrypio s , tačiau šios reikšmės sužinoti negalima, kol duomenys nesurinkti. Šią reikšmę galima gauti arba remiantis ekspertų nuomone, arba ankstesnių tyrimų rezultatais, arba gauti iš atrinktos kontrolinės imties. Paprastai kontrolinės imties atrinkimas yra per brangus, ir pasikliaujama kitais metodais. Tarkime, kad remiantis turima informacija, s apytiksliai yra 220 Lt, tada pagal formulę sužinoma:

$$\sqrt{n} = 1,96 \frac{220}{10} = 43,12 \approx 43 \text{ ir } n=1849.$$

Žinoma, imties neturi tiksliai sudaryti 1849 elementai, tačiau šis rezultatas – tai pageidautinas imties dydis. Kadangi vieno tyrimo metu paprastai tiriama daugiau nei vienas požymis, realiai būna sudėtingiau.

2 pavyzdys. Reikia nustatyti imties dydį tada, kai norima patikrinti, ar tikrai 35 proc. asmenų padaro nusikaltimus būdami neblaivūs.

Šiuo atveju dispersija:

$$s^2 = p(1 - p) = 0,35 \cdot (1 - 0,35) \approx 0,23.$$

Jeigu tenkina ± 5 proc. paklaida ir 68,3 proc. pasikliovimo lygmuo (kartojant tyrimą, tik 683 vienetai iš tūkstančio patenka į nurodytas imties paklaidos ribas), tai:

$$0,05 = 1 \cdot \frac{s}{\sqrt{n}} \text{ arba } n = \frac{s^2}{0,05^2}.$$

Šiuo atveju pageidautinas imties dydis yra:

$$n = \frac{0,23}{0,05^2} = 92 \text{ elementai.}$$

Jeigu manoma, kad 68,3 proc. pasiklovimo lygmuo nėra labai patikimas, pasirenkama 95 proc. pasiklovimo lygmenį. Šiuo atveju pageidautinas imties dydis yra:

$$0,05 = 1,96 \cdot \frac{s}{\sqrt{n}} \text{ arba } n = \frac{0,23 \cdot 1,96^2}{0,05^2} \approx 354 \text{ elementai.}$$

3 pavyzdys. Kai populiacijos dydis žinomas, imties dydžiui nustatyti galima taikyti šią formulę [Schwarze, 1993]:

$$n = \frac{N \cdot 1,96^2 \cdot p \cdot q}{\varepsilon^2 \cdot (N - 1) + 1,96^2 \cdot p \cdot q};$$

čia: N – populiacijos dydis;

Reikšmė 1,96 atitinka standartizuoto normaliojo skirstinio 95 proc. pasiklovimo lygmenį.

p yra numatoma įvykio baigmės tikimybė, kad nagrinėjamas požymis pasireikš tiriamoje populiacijoje (dažniausiai imama blogiausio varianto tikimybė – požymis būdingas pusei, t. y. 50 proc. populiacijos, ir pasirenkama $p=0,5$);

q yra tikimybė, kad nagrinėjamas požymis nepasireikš tiriamoje populiacijoje ($q=1-p=0,5$);

ε yra pageidautinas tikslumas, dažniausiai $\varepsilon=0,05$.

Jeigu populiacijos dydis yra, pavyzdžiui: $N=1000$ asmenų, tai pagal šią formulę skaičiuojant imties dydį $n=278$.

Kartais tenka spręsti ir kito tipo uždavinį – nustatyti paklaidos dydį, kai imtis žinoma. Pavyzdžiui: žinoma, kad populiacijos dydis $N=750$, o apklausti 183 asmenys. Šiuo atveju apskaičiuojamas paklaidos ε dydis, kuris apytiksliai yra $\varepsilon=0,063$, arba 6,3 proc. Vadinasi, nustačius savybę, būdingą 80 proc. apklaustųjų su 95 proc. patikimumu, galima sakyti,

kad šia savybe pasižymi nuo 73,7 proc. iki 86,3 proc. tiriamos populiacijos.

Pastaba. Imties dydžiui skaičiuoti internete galima rasti įvairių programinių skaičiuoklių.

1.8. Apklaustos taisyklės, apklaustos anketų pavyzdžiai

„Žmonės jaudina ne patys įvykiai, o jų nuomonė apie šiuos įvykius.“

Epiktetas

Apklaustos taisyklės

Apklausų metodistai teigia, kad apklausa iš tiesų yra menas, o gaudamos apklausų metu informacijos kokybė itin priklauso nuo tyrėjo. Visgi apklaustos metodologinės žinios ir įgūdžiai yra būtini, norint tinkamai atlikti apklausą ar pagerinti jau įgytą patirtį. Gaunamų atsakymų kokybė yra apklausai parinktų klausimų ir atskirų klausimų struktūros funkcija. Kiekvienas praktikuojantis tyrėjas žino kad nėra nei gerų, nei blogų klausimų, tik į skirtingus klausimus skirtingai atsakoma. Socialiniuose moksluose vyrauja klausimai, į kuriuos pateikiami nominalinio arba ranginio tipo atsakymai. Sudarant klausimus ir vykdant apklausą, reikia nepamiršti kelių paprastų taisyklių:

- Klausimai turi būti trumpi ir tikslūs, su kiek įmanoma mažiau dviprasmiškumo. Klausimas turi būti sudarytas taip, kad jį būtų įmanoma interpretuoti tik vienu būdu. Patartina vengti įterptinių frazių. Respondento atsakymo kokybė labai priklauso nuo to, kaip aiškiai respondentas supranta apklaustos tikslus. Klausimas, kuriame įterpta papildoma frazė, pavyzdžiui: „Kaip jūs vertinate ryšius (turint galvoje, kaip žmonės sugyvena, kokius vaidmenis jie vaidina) tarp tėvų ir vaikų?“, gali supainioti respondentą, ir į klausimą jis gali neatsakyti.
- Vienu metu pateikiamas tik vienas klausimas, nes į klausimus keliomis temomis retai kada išsamiai atsakoma. Pavyzdžiui, klausimai

mą „Kaip jūs vertinate dabartinės valdžios ir opozicijos atliekamą darbą?“ reiktų išskirti į du klausimus.

- Jei apklausa vykdoma tiesiogiai, vienu metu pateikiamas tik vienas klausimas. Pateikus klausimą, respondentui reikia duoti laiko apmąstymui. Tyloje nejaukiai pasijutęs tyrėjas perfrazuoja sakinį ir perfrazuodamas dažnai iš esmės užduoda kitą klausimą. Uždavus ir perfrazavus klausimą, būna sunku suprasti, į kurį klausimą respondentas atsakė.
- Patartina vengti klausimų, į kuriuos atsakymas yra iš anksto pateikiamas ar numatomas. Apklausos tikslas yra ne suformuoti respondento nuomonę ar įsitikinimus, o įvertinti respondento požiūrį. Tyrėjai turi būti labai atidūs, kad klausimais neperteiktų respondentui savo požiūrio. Pavyzdžiui: pateikęs klausimą „Ar gyvenimo kokybė yra svarbesnė už gyvenimo trukmę?“, tyrėjas tarsi perduoda respondentui numanomą žinią, kad tyrėjas gyvenimo kokybę vertina labiau už gyvenimo trukmę. Be to, jei tik įmanoma, patartina vengti dichotominių klausimų, kadangi atsakymuose „taip“ ir „ne“ yra labai nedaug informacijos.
- Reikia išmanyti respondento kalbos stilių (t. y. tarmę, kalbos savytumą, žargoną). Tyrėjas turi pritaikyti savo klausimus ir klausimų struktūrą taip, kad jie atitiktų respondento kalbos stilių, kuris nebūtinai yra toks pat, kaip ir tyrėjo.
- Patartina vengti klausimo „kodėl“. Klausimu „kodėl“ prašoma respondentų patvirtinti ankstesnius atsakymus, mintis ar nuomones, ir respondentai gali juos suprasti kaip grasinančius. Kai grasinama, respondentai atsako labai atsargiai. Tai savo ruožtu turi įtakos gaunamų duomenų kokybei. Taip tyrėjas rizikuoja prarasti vertingą informacijos dalį. Klausimas „kodėl“, pavyzdžiui: „Kodėl jūs norite išvažiuoti gyventi į kitą šalį?“ gali būti lengvai perfrazuotas: „Dėl kokių priežasčių jūs norite išvažiuoti gyventi į kitą šalį?“.
- Atsakymų kokybei įtakos gali turėti formuluojamo klausimo griežtumas. Pavyzdžiui, jei į klausimą „Ar jūs manote, kad reiktų leisti...?“ daugelis respondentų atsakys, kad nereikia leisti, tai į klausimą „Ar jūs manote, kad reikia uždrausti...?“, teigiamai atsakiusiųjų dažniausiai būna gerokai mažiau.

- Į klausimus numatant ranginio tipo atsakymus („puikiai“, „labai gerai“, „gerai“, „patenkinamai“, „nepatenkinamai“), turėtų būti ne mažiau kaip keturios pozicijos.

Apklaustos anketų pavyzdžiai

Sudarant apklaustos anketas, visada iškyla daug klausimų ir problemų. Liaudies išmintis teigia, kad geriau vieną kartą pamatyti nei šimtą kartų išgirsti. Todėl šiame skyrelyje pateikiama apklaustos anketų, kurias savo tyrimuose naudojo viešosios nuomonės ir rinkos tyrimų kompanija „TNS GALLUP“. Pirmoji viešosios nuomonės apklaustos tyrimų anketa buvo naudojama prieš Lietuvai stojant į Europos Sąjungą, antroji ir trečioji anketos – vartotojų nuomonės apklausai.

1 anketa

Narystė Europos Sąjungoje

Q1 Kiek Jums metų?

.....metų

Q2 Koks Jūsų išsilavinimas?

Pradinis	1
Vidurinis	2
Profesinis vidurinis (PTM).....	3
Aukštesnysis	4
Aukštasis	5

Q3 Jūsų užsiėmimas...

Dirbu nekvalifikuotą darbą	1
Dirbu kvalifikacijos reikalaujantį darbą	2
Pats sau darbdavys	3
Ūkininkas	4
Studentas	5
Namų šeimininkė	6
Pensininkas	7
Bedarbis	8

*PILDO INTERVIUOTOJAS:***Q4 Lytis:**

Vyras	1
Moteris	2

Q5 Vietovės dydis:

1 – 5 000 gyventojų	1
5 001 – 20 000 gyventojų	2
20 001 – 100 000 gyventojų	3
100 001 ir daugiau gyventojų	4

Q6 Šalis:

Bulgarija	1
Čekija	2
Estija	3
Vengrija	4
Latvija	5
Lietuva	6
Lenkija	7
Rumunija	8
Slovakija	9
Slovėnija	10
Turkija	11

EU 1. Prašau man pasakyti, jei Jūs dalyvautumėte referendume dėl Lietuvos narystės ES, balsuotumėte už ar prieš narystę?

Už	1
Prieš	2
Kitaip	3

EU 2. Yra manoma, kad Lietuva, įstojusi į ES, būtų antrarūšė šalis tarp kitų, turtingesnių jos narių. Ar Jūs sutinkate su tokia nuomone?

Visiškai sutinku	1
Sutinku	2
Nesutinku	3
Visiškai nesutinku	4
Kitaip.	5

EU 3. Jūsų manymu, ar Jūs asmeniškai arba Jūsų artimieji laimėtumėte ar praloštumėte šiose srityse, jei Lietuva taptų ES nare?

	BE ABEJONĖS LAIMĖTUME	GREIČIAU LAIMĖTUME NEI NE	GREIČIAU PRALOŠTU- ME NEI LAI- MĖTUME	BE ABEJONĖS PRALOŠ- TUME	DAR KITAIP	NĖRA AIŠKU
	1	2	3	4	5	6
a)	Išsilavinimo galimybės		1 2 3 4 5 6			
b)	Keliavimo galimybės		1 2 3 4 5 6			
c)	Darbo susiradimo galimybės		1 2 3 4 5 6			
d)	Prekių kainų lygis		1 2 3 4 5 6			
e)	Aplinka Jūsų gyvenamojoje vietoje		1 2 3 4 5 6			
f)	Pragyvenimo lygis		1 2 3 4 5 6			
g)	Pasididžiavimas savo šalimi .		1 2 3 4 5 6			

EU 4. Ar norėtumėte įsidarbinti kitoje ES šalyje po to, kai Lietuva taps Europos Sąjungos nare?

Tikrai taip	1
Greičiausiai taip	2
Greičiausiai ne	3
Tikrai ne	4
Kitaip.....	5

EU5 a) Prašau pasakyti, kuri šalis iš šio sąrašo Jums būtų pati patraukliausia dirbti?

ĮRAŠYKITE PIRMĄ PAGAL PATRAUKLUMĄ

.....

EU5 b) O kuri būtų antra pagal patrauklumą?

ĮRAŠYKITE

.....

EU5 c) Kuri šalis būtų mažiausiai patraukli dirbti?
ĮRAŠYKITE

.....

Q EU5 a-c	
Austrija	01
Belgija	02
Bulgarija	03
Kipras	04
Danija	05
Estija	06
Suomija	07
Prancūzija	08
Vokietija	09
Graikija	10
Vengrija	11
Airija	12
Italija	13
Latvija	14
Liuksemburgas	16
Malta	17
Lenkija	18
Portugalija	19
Rumunija	20
Slovakija	21
Slovėnija	22
Ispanija	23
Švedija	24
Čekija	25
Olandija	26
Turkija	27
Didžioji Britanija	28

Valstybėse su išvystyta rinkos ekonomika dažnai atliekami vartotojų nuomonės tyrimai. Šių tyrimų tikslas – sužinoti gyventojų nuomonę

apie siūlomas prekes ir paslaugas. Vartotojų nuomonės tyrimams naudojamų anketų struktūra yra paprastesnė, o pačios anketos neilgos. Susipažinkite su anketomis vartotojų nuomonei apie kavą ir alų sužinoti.

2 anketa

Kava

K1. Kaip dažnai per pastaruosius 6 mėnesius Jūs pirkote kavą?

1. Kartą per savaitę ar dažniau
2. 2–3 kartus per mėnesį
3. Kartą per mėnesį
4. Rečiau nei kartą per mėnesį
5. Pirkau seniau nei prieš 6 mėnesius
6. Visai nepirkau.

Kokias kavos rūšis žinote?

K2. *Pirma paminėta kavos rūšis:* _____

K3. *Kitos paminėtos kavos rūšys:* _____

K4. Kalbant apie kavą, su kuriuo teiginiu Jūs labiausiai sutinkate?

1. Yra viena kavos rūšis, tikrai geresnė už kitas
2. Yra keletas kavos rūšių, kurias galima laikyti geriausiomis
3. Visos kavos rūšys vienodos

K5. Kokių rūšių kavą (iš išvardytų kortelėje) Jūs pats pirkote per paskutiniuosius 6 mėnesius?

Atsakymus rašykite stulpelyje „Pirko per paskutiniuosius 6 mėnesius“

K6. Kokių rūšių kavą Jūs pirsite ateityje?

Atsakymus rašykite stulpelyje „Pirks ateityje“.

	K5 Pirko per paskutiniuosius 6 mėnesius	K6 Pirks ateityje
Paulig	1	1
Nescafe	2	2
Jacobs	3	3
Edusho	4	4
Folgers	5	5
Elite	6	6
Merrild	7	7
Ar Mani	8	8
Tchibo	9	9
Luxus	10	10
Kita	11	11

Kaip Jums patinka išvardytų rūšių kava? Įvertinkite dešimties balų skalėje, kur 1 balas reiškia „visai nepatinka“ ir 10 balų reiškia „labai patinka“.

	Visai nepatinka										Labai patinka	<i>Sunku pasakyti</i>
K7. Paulig	1	2	3	4	5	6	7	8	9	10	11	
K8. Nescafe	1	2	3	4	5	6	7	8	9	10	11	
K9. Jacobs	1	2	3	4	5	6	7	8	9	10	11	
K10. Edusho	1	2	3	4	5	6	7	8	9	10	11	
K11. Folgers	1	2	3	4	5	6	7	8	9	10	11	
K12. Merrild	1	2	3	4	5	6	7	8	9	10	11	
K13. Elite	1	2	3	4	5	6	7	8	9	10	11	
K14. Ar Mani	1	2	3	4	5	6	7	8	9	10	11	
K15. Tchibo	1	2	3	4	5	6	7	8	9	10	11	
K16. Luxus	1	2	3	4	5	6	7	8	9	10	11	

3 anketa**Alus****A1. Kaip dažnai per pastaruosius 6 mėnesius Jūs pirkote alų?**

1. Kartą per savaitę ar dažniau
2. 2–3 kartus per mėnesį
3. Kartą per mėnesį
4. Rečiau nei kartą per mėnesį
5. Pirko seniau nei prieš 6 mėnesius ⇒ *pereiti prie klausimo K1.*
6. Visai nepirko ⇒ *pereiti prie klausimo K1.*

Kokias alaus rūšis žinote? *Nerodyti alaus rūšių sąrašo.*

A2. *Pirma paminėta alaus rūšis:* _____

A3. *Kitos paminėtos alaus rūšys:* _____

A4. Kalbant apie alų, su kuriuo teiginiu Jūs labiausiai sutinkate?
Rodyti P2 KORTELE.

1. Yra viena alaus rūšis, tikrai geresnė už kitas
2. Yra keletas alaus rūšių, kurias galima laikyti geriausiomis
3. Visos alaus rūšys vienodos

Paduokite respondentui P3 KORTELE.

A5. Kokių rūšių alų (iš išvardytų kortelėje) Jūs pats pirkote per paskutiniuosius 6 mėnesius?

Atsakymus rašykite stulpelyje „Pirko per paskutiniuosius 6 mėnesius“

A6. Kokių rūšių alų Jūs pirksite ateityje?

Atsakymus rašykite stulpelyje „Pirks ateityje“.

	A5 Pirko per paskutiniuosius 6 mėnesius	A6 Pirks ateityje
Kalnapilio	1	1
Utenos	2	2
Švyturio	3	3
Ragučio	4	4
Vilniaus Tauro	5	5
Gubernijos	6	6
Faxe	7	7
Carlsberg	8	8
Kita	9	9

Paduokite respondentui P4 KORTELE.

Kaip Jums patinka išvardytų rūšių alus? Įvertinkite dešimties balų skalėje, kur 1 balas reiškia „visai nepatinka“ ir 10 balų reiškia „labai patinka“.

	Visai nepatinka										Labai patinka	<i>Sunku pasakyti</i>
A7. Kalnapilio	1	2	3	4	5	6	7	8	9	10	11	
A8. Utenos	1	2	3	4	5	6	7	8	9	10	11	
A9. Švyturio	1	2	3	4	5	6	7	8	9	10	11	
A10. Ragučio	1	2	3	4	5	6	7	8	9	10	11	
A11. Vilniaus Tauro	1	2	3	4	5	6	7	8	9	10	11	
A12. Gubernijos	1	2	3	4	5	6	7	8	9	10	11	
A13. Carlsberg	1	2	3	4	5	6	7	8	9	10	11	

Kontroliniai klausimai

- 1.1. Kas tai yra populiacija ir imtis? Ar visada populiacija ir imtis skiriasi?
- 1.2. Kaip skirstomi tyrėjai?
- 1.3. Kokia seka iš duomenų gaunamos žinios?
- 1.4. Kokie yra pagrindiniai statistinio tyrimo metodai?
- 1.5. Kokios pagrindinės apklausos formos?
- 1.6. Kokios dalys išskiriamos anketoje (apklausos lape)?
- 1.7. Kaip skirstomi anketos pagrindinės dalies klausimai?
- 1.8. Kokie duomenys yra privačių duomenų klasės?
- 1.9. Koks yra oficialios tyrimo ataskaitos turinys?
- 1.10. Kokie yra pagrindiniai imčių sudarymo būdai?
- 1.11. Nuo ko priklauso atsakymų matavimo klaidos?
- 1.12. Nuo ko priklauso imties dydis?
- 1.13. Kokios yra duomenų matavimo skalės? Pateikite pavyzdžių.



2. PIRMINĖ DUOMENŲ APŽVALGA arba APRAŠOMOJI STATISTIKA

Prieš atliekant nuodugnesnę surinktų duomenų statistinę analizę, pirmiausia reikia nagrinėti kiekvieną kintamąjį (požymį) *atskirai*. Kai kiekvienas kintamasis nagrinėjamas atskirai, statistinė analizė yra vienu matė. Šiai analizei atlikti naudojami aprašomosios statistikos metodai. Aprašomosios statistikos metodus sudaro duomenų grupavimas, dažnių lentelės, statistinių charakteristikų (duomenų padėties ir sklaidos charakteristikų) skaičiavimas ir grafinis stebėjimų vaizdavimas. Sudaromos kintamųjų dažnių lentelės, kuriose pateikiami nominalinėje arba (esant nedideliam kategorijų skaičiui) rangų skalėje išmatuotų kintamųjų dažniai.

Pradinei duomenų analizei, be vienmačių dažnių lentelių, dar sudaromos porinės (ir daugiau kintamųjų) dažnių lentelės. Porinės dažnių lentelės atskleidžia tarp kintamųjų egzistuojančius ryšius. Tokios dažnių lentelės paprastai sudaromos nominaliniams ir ranginiams kintamiesiems su santykinai mažu stebimų reikšmių (kategorijų) skaičiumi. Porinėms lentelėms sudaryti *ne naudojami* kiekybiniai kintamieji. Jeigu į tokią lentelę būtina įtraukti kiekybinį kintamąjį (pajamas ir pan.), tai pirmiausia šio kintamojo reikšmes reikia suskirstyti į keletą intervalų (mažos pajamos, vidutinės, didelės, labai didelės).

Aprašomoji statistika, be dažnių lentelių ir diagramų, nominaliniams kintamiesiems jokių kitokių metodų nesūlo. Ranginiams kintamiesiems ir kiekybiniam kintamiesiems, išmatuotiems intervalų ar santykių skalėse, bet neturintiems normaliojo skirstinio, šalia dažnių lentelių ir diagramų dažniausiai dar skaičiuojama vidutinė (centrinė) atsitiktinio dydžio reikšmė (mediana) ir kvartiliai. Kvartiliais vadinami skaičiai, dalinantys variacinę eilutę (variacinė eilutė – išdėstyta nemažėjimo tvarka duomenų eilutė) į keturias maždaug lygias dalis. Kintamieji, išmatuoti intervalų ar santykių skalėse ir turintys normalųjį skirstinį, grupuojami į intervalus (tam reikia nustatyti grupavimo intervalų skaičių, jų plotį ir intervalų kraštinius taškus), skaičiuojami dažniai, piešiamos diagramos (histogramos, linijinės diagramos ir pan.) ir skaičiuojamos centro ir sklaidos charakteristikos.

2.1. Dažnių analizė

Praktiškai atliekant tyrimą, pirmoji pažintis su duomenimis prasideda nuo dažnių lentelių. Vienmatė dažnių lentelė yra paprasčiausia pavadinimų (nominaliųjų) ir ranginių kintamųjų analizės forma. Dažnių lentele naudojamosi tiriant, kaip atskiro kintamojo įgyjamos reikšmės pasiskirstę imtyje. Pavyzdžiui: socialiniuose tyrimuose iš dažnių lentelių matyti, kiek tyrime dalyvavo vyrų, kiek moterų, kokia buvo apklausiųjų šeimyninė padėtis, koks jų išsilavinimas.

STATISTICA programos 6.0 versijoje pagrindiniai aprašomosios statistikos metodai pateikiami komandų, parametrų, statusų sąrašo (meniū) skyriaus *Statistics* dalyje *Basic Statistics/Tables*:

Statistics

Basic Statistics/Tables

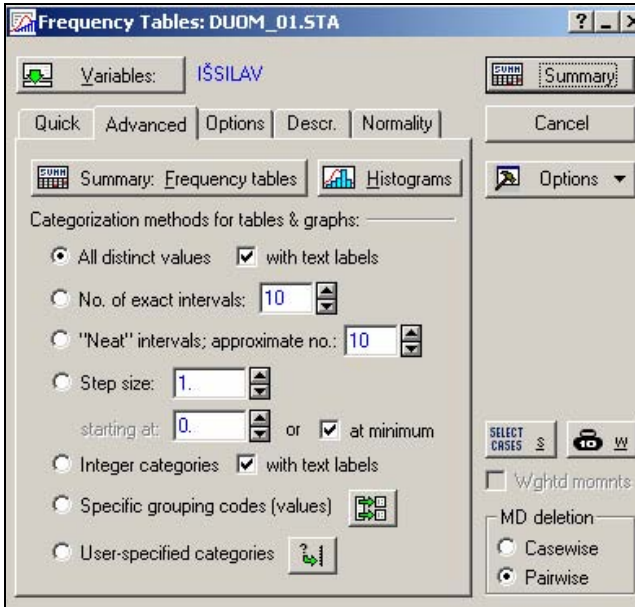
Frequency tables

Gauname dialoginį langą *Frequency Tables* (2.1 pav.), kuriuo remiantis galima apskaičiuoti dažnių lenteles, nubraižyti histogramas ir rasti aprašomąsias statistikas.

Frequency Tables dialoginiame lange yra daug įvairių pasirinkimo tipų. Naudojant *Advanced* formą, kaip bus apskaičiuojama dažnių lentelė priklauso nuo to, koks skaičiavimo metodas nurodytas skyriuje *Categorization methods for tables & graphs*. Atkreipiame dėmesį į tai kad:

- jeigu nurodomas metodas *All distinct values* (*Visos skirtingos reikšmės*), tai bus skaičiuojami pasirinkto kintamojo visų skirtingų reikšmių dažniai. Jei, be to, dar nurodoma *with text values*, tai skirtumai bus nustatomi pagal skirtingas teksto reikšmes.
- jeigu nurodomas metodas *No of exact intervals* (*Tikslus intervalų skaičius*), tai visas reikšmių kitimo diapazonas bus padalintas į nurodytą skaičių intervalų.
- jeigu parenkamas skaičiavimo metodas *“Neat” intervals; app. no* (*approximate number*) tai dažnių lentelių viršutinis ir apatinis režiai ir kitimo intervalai bus „nušlifuoti“, t. y. suapvalinti iki artimiausio sveiką skaičiaus. Kartais taip sudaryta lentelė yra lengviau suprantama ir patogesnė.
- jeigu nurodomas metodas *Step size* (*Intervalo ilgis*), tai skaičiuojant dažnių lentelę ir braižant histogramą bus laikomasi vartotojo

nurodyto padalinimo intervalo ilgio. Čia galima nurodyti ir minimalią reikšmę (*starting at*). Nurodžius minimalią reikšmę, visos už ją mažesnės reikšmės bus ignoruojamos.



2.1 pav. STATISTICA programos dažnių lentelių dialoginis langas

- parinkus metodą *Integer categories* (*Sveikos reikšmės*), dažnių lentelės (ir histogramos) bus skaičiuojamos, pradedant nuo mažiausios sveikosios nagrinėjamo kintamojo reikšmės. Visos realios reikšmės šiuo atveju yra ignoruojamos. Nurodžius *with text values*, dažnių lentelių ir histogramų bus ne sveikosios reikšmės, o teksto reikšmės (*vyr'ai, moterys*).
- jeigu nurodomas metodas *Specific grouping codes* (*Konkretūs grupių kodai*), tai dažnių lentelė bus sudaroma pagal vartotojo nurodytus kodus. Nenurodyti kodai į skaičiavimą neįtraukiami. Nenurodžius jokių kodų, skaičiavimams naudojamos visos sveikosios reikšmės.

- metodas *User-specified categories* (Vartotojo nurodytos reikšmės) sudaro galimybę užrašyti sąlygas, pagal kurias atrenkami duomenys analizei. Tokių sąlygų gali būti iki 16. Sąlygų išraiška gali būti sudėtinga, jas gali sudaryti keletas kintamųjų. Sąlygos analizuojamos eilės tvarka.

Dažnių analizės rezultatai

Jeigu dažnių lentelės sudaromos kiekvienam požymiui atskirai, STATISTICA programoje dažnių lentelė gaunama nurodžius kintamojo vardą (klavišu *Variables*), parinkus skaičiavimo metodą ir paspaudus klavišą *Frequency tables*. Jeigu nurodoma kelių kintamųjų vardai, jiems bus sudarytos atskiros lentelės. Pavyzdžiui: požymiui *Išsilavinimas* gauname tokio tipo dažnių lentelę (2.1 lentelė):

2.1 lentelė. Išsilavinimas

IŠSILAV				
	Count	Cumul. Count	Percent	Cumul. Percent
Aukštasis	252	252	41.31	41.31
Vidurinis	318	570	52.13	93.44
Nebaigtas vidurinis	25	595	4.10	97.54
Pradinis	10	605	1.64	99.18
Missing	5	610	0.82	100.00

Eilutė *Missing* – tai stebėjimai, kuriems išsilavinimas nebuvo nurodytas. Šiuo atveju penki respondentai nenurodė savo išsilavinimo. Pirmaime stulpelyje išvardijami visi galimi požymio atsakymai (kategorijos). Stulpelyje *Count* (*Dažniai*) apskaičiuota, kiek yra kiekvieno tipo atsakymų, t. y. kiek asmenų turi aukštąjį išsilavinimą, kiek vidurinį ir t.t. Stulpelyje *Cumul. Count* (*Sukauptieji dažniai*) dažniai yra sudedami; iš šio stulpelio reikšmių matyti, pavyzdžiui: kiek asmenų turi aukštąjį ir vidurinį išsilavinimą (570) arba aukštąjį, vidurinį ir nebaigtą vidurinį išsilavinimą (595) ir t.t. Čia pateiktas ir bendras asmenų skaičius – 610. Stulpelyje *Percent* (*Procentai*) apskaičiuota, kiek procentų nuo bendro skaičiaus sudaro kiekvieno tipo atsakymai, o *Cumul Percent* (*Sukauptieji dažniai procentais*) pateiktos procentinės sukauptųjų dažnių reikšmės.

Porinės dažnių lentelės

Paprasčiausia porinės dažnių lentelės forma yra 2×2 lentelė, gaunama tuo atveju, kai tiriama du kintamieji, kurių kiekvienas gali įgyti tik dvi reikšmes (dichotominiai kintamieji). Pavyzdžiui, atliekant tyrimą „Smurtas prieš moteris“, mieste ir kaime gyvenančios ir smurtą patyrusios moterys buvo paprašytos nurodyti, ar jos kreipėsi pagalbos į atitinkamas institucijas. Apklauskos duomenų bylos (failo) fragmentas atrodo taip (2.2 lentelė).

2.2 lentelė. Pradinių duomenų apie smurtą prieš moteris fragmentas

ID	Gyvenamoji vieta	Ar kreipėtės pagalbos į atitinkamas institucijas?
1	Miestas	Taip
2	Miestas	Ne
3	Kaimas	Ne
4	Kaimas	Taip
5	Miestas	Taip

Porinės dažnių lentelės skaičiavimo rezultatai gali būti teikiami kelių būdu, iš kurių dažniausiai naudojamas pateiktasis 2.3 lentelėje.

2.3 lentelė. Požymių „Gyvenamoji vieta“ ir „Ar kreipėsi pagalbos į atitinkamas institucijas?“ porinė dažnių lentelė

	Ar kreipėsi		Row
	Taip	Ne	Totals
Miestas	2	1	3
Total %	40.00	20.00	60.00
Kaimas	1	1	2
Total %	20.00	20.00	40.00
All Grps	3	2	5
Total %	60.00	40.00	

Iš 2.3 lentelės ląstelių aišku, kiek moterų, gyvenančių mieste, kreipėsi pagalbos į atitinkamas institucijas (2), kiek nesikreipė (1), kiek kaime gyvenančių moterų kreipėsi pagalbos (1), kiek jų nesikreipė (1).

Kartu su absoliutiniais dydžiais pateikta ir duomenų procentinė išraiška, kuri palengvina kintamuosius palyginti. Apatinėje eilutėje ir kraštiniame dešiniajame stulpelyje pateikti suminiai arba kraštiniai (*marginal*) dažniai, kurie padeda įvertinti dažnius, pateiktus atskiruose stulpeliuose ar eilutėse. Iš kraštiniame dešiniajame stulpelyje pateiktų rezultatų matyti, kad šio tyrimo fragmento 60 proc. apklaustųjų gyvena mieste, o 40 proc. – kaime. Iš apatinės eilutės matyti, kad 60 proc. kreipėsi pagalbos, o 40 proc. – nesikreipė.

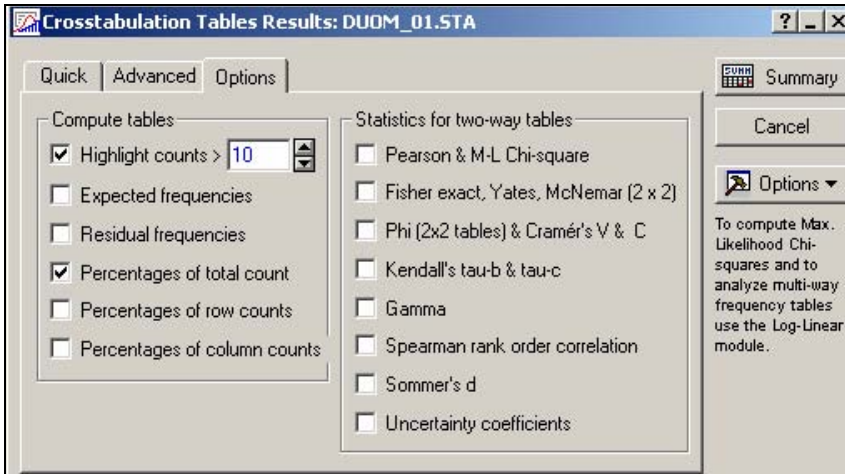
Norint **STATISTICA programa** atlikti porinę ir kitokią kintamųjų dažnių analizę, renkamos šios komandos:

Statistics

Basic Statistics/Tables

Tables and banners

Gauname dialoginį langą *Crosstabulation Tables*, kuriame nurodome ne mažiau kaip dvejų kintamųjų pavadinimus. Įvedus kintamųjų pavadinimus, parametrų parinkimo lange *Options* (2.2 pav.), galima pasirinkti pageidautiną dažnių lentelės pavidalą (ar bus pateikiami tikėtini dažniai, ar skaičiuojami procentai ir pan.) ir statistinius kriterijus ryšiams tarp požymių tikrinti. Kai kuriuos iš šių kriterijų nagrinėsime tolesniuose skyriuose.

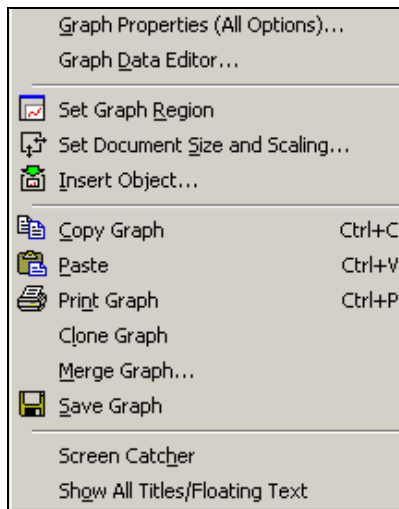


2.2 pav. *Crosstabulation Tables Results* dažnių lentelės parametrų parinkimo langas *Options*

2.2. Grafinis duomenų vaizdavimas

Įvairios diagramos, be abejonės, yra pats populiariausias duomenų vaizdavimo būdas. Informacija, pateikta bet kokio tipo lentelėje, visada tam tikru būdu gali būti pateikta grafiškai. Statistinės programos siūlo įvairiausių tipų diagramas. Standartiniai grafikai sudaromi naudojantis vartotojo aplinkos komandų, parametrų sąrašo (menu) procedūromis. Toks diagramų sudarymo būdas yra labai lengvas. Pasirinkus norimą diagramos tipą, nurodomas kintamojo vardas ir pagal iš anksto nustatytą schemą iš kintamojo reikšmių sudaroma diagrama. Sudarytą diagramą lengva redaguoti. Redagavimo principas daugelyje šiuolaikinių statistinių programų yra vienodas. Norint pakeisti kurį nors užrašą, liniją ar spalvą, tereikia du kartus spustelėti kairinį pelės klavišą ant norimo redaguoti elemento.

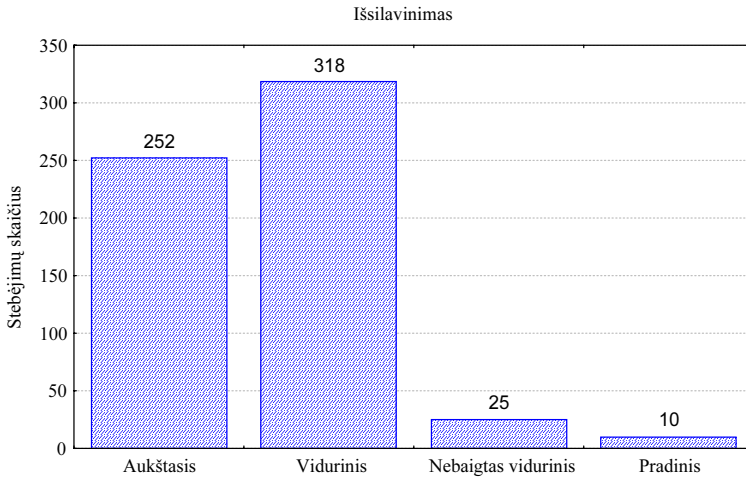
STATISTICA programa diagramoms redaguoti galima gauti papildomus langus. Ant diagramos paspaudus dešinįjį klavišą iškviečiamas papildomas komandų, parametrų sąrašas (menu) (2.3 pav.) diagramoms redaguoti.



2.3 pav. STATISTICA programa diagramų redagavimo komandų, parametrų sąrašas (menu), iškviečiamas dešiniu ju pelės klavišu

Dažnių lentelių grafinis pavidalas

Rezultatus, pateiktus dažnių lentelėje, galima pavaizduoti grafiškai. STATISTICA programoje tam skirtas dialoginio lango *Frequency Tables* klavišas *Histograms* (2.1 pav.). Paprasčiausiai dažniai vaizduojami stulpeliais, kurių viršuje kartais užrašomos apskaičiuotos dažnių reikšmės. Tokia dažnių diagrama yra vadinama stulpeline diagrama arba histograma (2.4 pav.).



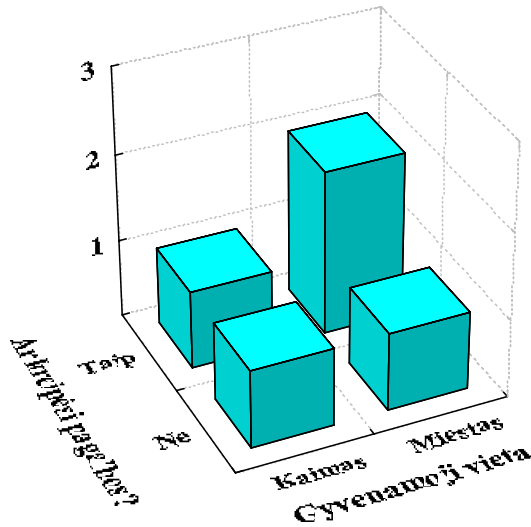
2.4 pav. Dažnių lentelės histograma

Porinėms dažnių lentelėms vaizduoti STATISTICA programoje yra įvairių stulpelinių diagramų. Kiek įdomesnė yra trimatė diagrama (2.5 pav.), gaunama parinkus komandų, parametrų sąrašo (menu) skyrių *Graphs*:

Graphs

3D Sequential Graphs

Bivariate Histograms



2.5 pav. Porinės dažnių lentelės trimatė diagrama

Ši grafiką galima sukoti ir parinkti tinkamiausią jo pavidalą. Sukimo komanda iškviečiama ant diagramos paspaudus dešinįjį pelės klavišą ir pasirodžiusiame komandų, parametrų sąrašė (menu) pasirinkus pirmąją eilutę *Graph Properties (All Options)* (2.3 pav.). Pateiktame komandų sąrašė parenkamas sąrašo punktas

Point of View

Analytic exploratory spin options

Ši diagrama, pavadinimu *3D Histograms*, pateikiama įvairiuose duomenų analizės languose.

2.3. Duomenų padėties ir sklaidos charakteristikos

Kintamiesiems, išmatuotiems intervalų ar santykių skalėmis ir turintiems normalųjį skirstinį, dažniausiai skaičiuojamas vidurkis ir standartinis nuokrypis. Kartais skaičiuojamas vidurkio standartinis nuokry-

pis (standartinė klaida). Statistinės programos pateikia daug galimų duomenų padėties ir sklaidos charakteristikos skaičiavimo pasirinkimo galimybių (2.6 pav.). Tačiau pateikiant skaičiavimo rezultatus patartina apsiriboti viena, geriausiai duomenims tinkančia, centro charakteristika (vidurkiu arba mediana) ir viena sklaidos charakteristika (dažniausiai pateikiamas standartinis nuokrypis arba standartinė klaida).

STATISTICA programoje duomenų padėties ir sklaidos charakteristikos skaičiuojamos komandų, parametru, statusų sąrašo (menu) skyriuje *Statistics*:

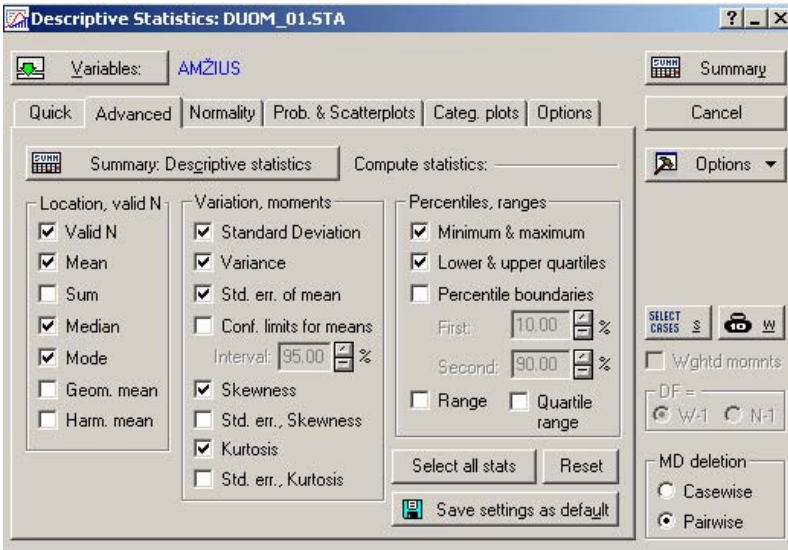
Statistics

Basic Statistics/Tables

Descriptive statistics

Descriptive statistics komandos pateiktos keliuose languose. Lange *Options* reikia atkreipti dėmesį į keletą pasirinkimo variantų.

Pažymėjus *Display long variable names*, rezultatų lange kartu su trumpaisiais kintamųjų vardais bus parodyti ir ilgieji kintamųjų vardai (jei jie nurodyti).



2.6 pav. STATISTICA programos duomenų charakteristikų pasirinkimo (*Advanced*) langas

Pasirinkimas *Extend precision calculations* sudaro galimybę padidinti skaičiavimo tikslumą. Šis pasirinkimas turėtų būti naudojamas, kai gaunama labai maža duomenų sklaidos charakteristikos – dispersijos – reikšmė.

Pagrindinis šio punkto langas – *Advanced*, kuriame pateikiamas galimų duomenų padėties ir sklaidos charakteristikų sąrašas (2.6 pav.).

Šiame lange pateikiamos keturios charakteristikų grupės.

Pirmąją grupę sudaro šios duomenų padėties (centro) charakteristikos:

- *Vidurkis* arba aritmetinis vidurkis (*Mean*) – tai populiariausias duomenų vidurio (centro) matas. Jis apskaičiuojamas sudėjus visas reikšmes ir padalinus iš reikšmių skaičius. Vidurkis, apskaičiuotas iš imties, yra empirinis ir žymimas \bar{x} . Skirtingai nei empirinis vidurkis, populiacijos vidurkis žymimas raide μ .
- *Mediana* (*Median*) – tai matavimų skalės taškas, kuris stebimas reikšmes padalina į dvi lygias dalis. Pavyzdžiui, jei išmatavę gavome reikšmes

7 1 8 8 9 1 6 8 5

tai pirmiausia jas reikia išrikiuoti didėjimo tvarka, t. y. sudaryti variacinę eilutę:

1 1 5 6 7 8 8 8 9

ir surasti vidurinę reikšmę (šiuo atveju 7). Jei stebėjimų skaičius yra lyginis, tai surandamas dviejų vidurinių reikšmių aritmetinis vidurkis.

- *Moda* (*Mode*) – tai dažniausiai pasikartojanti reikšmė. Jeigu kelios reikšmės kartojasi vienodai kartų, tai pasirenkama mažiausia reikšmė.

Antrąją grupę sudaro šios duomenų sklaidos charakteristikos:

- *Standartinis nuokrypis* (*Standard Deviation*) – tai matuojamų dydžių sklaidos apie vidurkį matas. Jis apskaičiuojamas ištraukiant kvadratinę šaknį iš dispersijos. Imties (empirinis) standartinis nuokrypis žymimas raide s , o populiacijos standartinis nuokrypis – raide σ . Jei duomenų skirstinys yra normalusis, tai į intervalą, kuris gaunamas atidėjus į abi puses nuo vidurkio po standartinį

nuokrypį s ($\bar{x} - s, \bar{x} + s$), patenka apytiksliai 68 proc. visų duomenų. Į intervalą ($\bar{x} - 2s, \bar{x} + 2s$) patenka apytiksliai 95,4 proc. visų duomenų, o į intervalą ($\bar{x} - 3s, \bar{x} + 3s$) patenka beveik visi duomenys (tiksliau – 99,7 proc.).

- *Dispersija (Variance)* apskaičiuojama sudėjus kiekvienos reikšmės atstumo (skirtumo) nuo vidurkio kvadratus ir gautą sumą padalijus iš visų reikšmių skaičiaus be vieneto ($n-1$). Dispersija, kaip ir standartinis nuokrypis, – tai duomenų sklaida apie vidurkį. Kadangi dispersija apskaičiuojama keliant reikšmės kvadratu, tai ir dispersijos matavimo vienetai keliami kvadratu. Standartinis nuokrypis, kuris yra šaknis iš dispersijos, matuojamas tokiais pačiais matavimo vienetais, kaip ir patys duomenys, todėl standartinį nuokrypį lengviau suprasti ir interpretuoti.
- *Vidurkio standartinis nuokrypis arba standartinė klaida (Standard Error of Mean)* yra apskaičiuojamas imties standartinį nuokrypį s padalinus iš kvadratinės šaknies, ištrauktos iš visų duomenų skaičiaus (imties tūrio n). Vidurkio standartinis nuokrypis, kaip ir imties standartinis nuokrypis, padeda nustatyti atskirų reikšmių išsibarstymo intervalą. Su 99,7 proc. tikimybe galima teigti, kad populiacijos vidurkis yra intervale, kuris gaunamas atidėjus į abi puses nuo vidurkio po tris vidurkio standartinius nuokrypius ($\bar{x} - 3s, \bar{x} + 3s$); su 95,4 proc. tikimybe galima teigti, kad populiacijos vidurkis yra intervale, gautame atidėjus į abi puses nuo vidurkio po du vidurkio standartinius nuokrypius ($\bar{x} - 2s, \bar{x} + 2s$); o jeigu nuo vidurkio atidėsime po vieną vidurkio standartinį nuokrypį ($\bar{x} - s, \bar{x} + s$), tai tikimybė, kad į šį intervalą pateks populiacijos vidurkis yra apytiksliai 68 proc.
- *Vidurkio pasikliautinis intervalas (confidence limits of mean)*. Vidurkio standartinis nuokrypis arba standartinė klaida leidžia apskaičiuoti vidurkio pasikliautinąjį intervalą. Su 95,4 proc. tikimybe galima teigti, kad intervale, kurį sudaro du vidurkio standartiniai nuokrypiai, atidėti į abi puses nuo imties vidurkio, bus populiacijos vidurkis. Atidėję ne po 2, o po 1,96 vidurkio standartinius nuokrypius, gausime tikslią 95 proc. tikimybę. Šis intervalas ir yra 95 proc. vidurkio pasikliautinis intervalas.

Trečiąją grupę sudaro du asimetriškumo matai:

- *Asimetrijos koeficientu (Skewness)* matuojama, ar duomenų išsidėstymas skiriasi nuo simetrinio išsidėstymo, t. y. nuo tokio duomenų išsidėstymo, kai reikšmės išsidėsto vienodais atstumais į abi puses nuo vidurkio. Jeigu duomenys turi normalųjį skirstinį, tai asimetrijos koeficientas yra nulis. Asimetrijos koeficientą galima naudoti kaip vieną iš būdų duomenų normalumui tikrinti. Jeigu asimetrijos koeficientas iš esmės skiriasi nuo nulio, tai duomenų skirstinys nėra normalusis. Kai asimetrijos koeficientas didesnis už nulį, asimetrija teigiama, kai mažesnis, – neigiama.
- *Eksceso koeficientas (Kurtosis)* rodo, ar histograma yra aukšta su trumpomis „uodegomis“, ar, atvirkščiai, plokščia su ilgomis „uodegomis“. Eksceso koeficientas parinktas taip, kad tuo atveju, kai duomenų skirstinys yra normalusis, eksceso koeficientas yra nulis.

Į atskirą grupę (ketvirtąją) išskirti procentiliai:

Skaičiai, kurie suskirsto variacinę eilutę į 100 apytiksliai lygių dalių, yra *procentiliai*, o skaičiai, dalinantys variacinę eilutę į keturias maždaug lygias dalis, – *kvartilai*. Dažniausiai kvartilai žymimi raidėmis Q_1, Q_2, Q_3 . Pirmasis kvartilis Q_1 – tai toks taškas išmatuotų reikšmių skalėje, nuo kurio kairėje pusėje lieka 25 proc. reikšmių. Antrasis kvartilis Q_2 – tai taškas, kurio kairėje pusėje yra 50 proc. reikšmių. Antrasis kvartilis sutampa su mediana. Trečiojo kvartilio Q_3 kairėje pusėje yra 75 proc. reikšmių.

Kaip viena iš galimų duomenų centro charakteristikų naudojamas pirmojo ir trečiojo kvartilių vidurkis $(Q_1 + Q_3)/2$.

Skaičiai, kurie duomenų eilutę dalija į 10 lygių dalių, yra *deciliai*.

Skaičiuojama pagal šias formules:

Aritmetinis vidurkis yra visų eilutės elementų suma, padalyta iš jų skaičiaus:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}. \quad (2.1)$$

Mediana yra vidurinė reikšmė. Ji nustatoma iš *didėjimo tvarka* surašytų skaičių (variacinės eilutės x_1, x_2, \dots, x_n), išrinkus vidurinį. Esant

lyginiam reikšmių skaičiui, dvi vidurinės reikšmės sudedamos ir dalijamos per pusę:

$$Md = \begin{cases} x_{(n+1)/2}, & \text{kai } n - \text{nelyginis} , \\ \frac{x_{n/2} + x_{n/2+1}}{2}, & \text{kai } n - \text{lyginis} . \end{cases} \quad (2.2)$$

Dispersija parodo duomenų sklaidą apie vidurkį:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}. \quad (2.3)$$

Standartinis nuokrypis yra dispersijos kvadratinė šaknis:

$$s = \sqrt{s^2}. \quad (2.4)$$

Variacijos (kitimo) koeficientas taikomas lyginant skirtingų duomenų aibių sklaidas:

$$CV = \frac{s}{\bar{x}}. \quad (2.5)$$

Duomenų aibės plotis: $R = x_{\max} - x_{\min}$.

Asimetrijos koeficientas:

$$As = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^3}{s^3}. \quad (2.6)$$

Eksceso koeficientas:

$$Ex = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^4}{s^4} - 3. \quad (2.7)$$

2.4. Nesudėtingos apklausos anketos pradinės duomenų analizės pavyzdys

Iliustruojant išdėstytas mintis, pateikiame nesudėtingą apklausos anketą ir pradinės duomenų analizės pavyzdį. Ši mokomąji tyrimą studijų metu atliko Mykolo Romerio universiteto I kurso studentai.

Viešosios nuomonės tyrimai *Tema: organų donorystė*

Tyrimo tikslas, imties apibūdinimas, atlikto tyrimo data

1. Šios anketos klausimais siekiama sužinoti žmonių požiūrį į organų donorystę bei išsiaiškinti, ar apskritai yra pakankamai žinių apie tai.
2. Tyrimas yra bandomasis, todėl statistinio tyrimo imtis yra maža – tik 30 žmonių.
3. Tyrimas atliktas 2004 m. kovo mėnesį.
4. Tiriama populiacija – pilnamečiai Lietuvos gyventojai. Tikimybiniė imtis sudaryta atsitiktinės atrankos metodu.
5. Atsakymus pateikė 100 procentų, t. y. visi 30 asmenų.

Apklausus atsitiktinai pasirinktus respondentus, galima daryti pradines išvadas apie žmonių požiūrį į organų donorystę ir analizuoti jų žinias.

Klausimynas:

1. Lytis
 - Vyras
 - Moteris
2. Gimimo metai:
3. Išsilavinimas:
 - Nebaigtas vidurinis
 - Vidurinis
 - Profesinis
 - Aukštesnysis
 - Nebaigtas aukštesnysis
 - Nebaigtas aukštasis
 - Aukštasis

4. Ar žinote apie tokį reiškinį, kaip organų donorystė?
 - Taip
 - Ne
5. Ar sutiktumėte paaukoti organus po savo mirties?
 - Taip
 - Ne
6. Ar sutinkate su teiginiu, jog donorystė labai reikalinga?
 - Visiškai sutinku
 - Labiau sutinku, nei prieštarauju
 - Nei sutinku, nei nesutinku
 - Labiau nesutinku, nei sutinku
 - Visiškai nesutinku
7. Ar žinote, kas yra donoro kortelė?
 - Taip, žinau
 - Kažką esu girdėjęs, bet tiksliai nežinau
 - Ne, nežinau
8. Ar aukotumėte kokį nors organą dar būdamas gyvas?
 - Be jokių abejonių
 - Aukočiau šeimai, giminaičiams ir artimiems draugams
 - Aukočiau šeimai bei giminaičiams
 - Aukočiau tik artimiausiems šeimos nariams
 - Tikrai neaukočiau
9. Kaip reaguojate/reaguotumėte į žmogų, planuojantį atiduoti savo organus po mirties?
 - Teigiamai
 - Neigiamai
 - Neturiu nuomonės
10. Ar artimiausiu metu planuojate įsigyti donoro kortelę?
 - Taip
 - Gal būt
 - Ne
 - Nesusimąščiau apie tai
11. Skalėje tarp 1–10 įvertinkite žmonių, laukiančių transplantacijos padėtį (1 – labai trūksta organų, 5 – organų pakanka, 10 – organų per daug). Skaičių apibraukite.
1 2 3 4 5 6 7 8 9 10

Apklauso rezultatai pateikti 2.4 lentelėje.

2.4 lentelė. Viešosios nuomonės apie organų donorystę tyrimo rezultatai

Lytis	Am-žius	Išsilavinimas	Ar žino, kas yra donorystė?	Ar sutiktų tapti donoru po mirties?	Ar reikalinga donorystė?	Ar žino, kas yra donoro kortelė?	Ar aukotų organus dar būdami gyvi?	Kaip reaguoja i ketinančius paaukoti organą?	Ar planuoja įsigyti donoro kortelę?	Kaip vertina žmonių, laukiančių transplantacijos, padėtį?
1	2	3	4	5	6	7	8	9	10	11
moteris	56	aukštesnysis	taip	taip	visiškai nesutinka	kažką žino	tik šeimai	neturi nuomonės	ne	2
vyras	34	aukštesnysis	taip	taip	visiškai sutinka	nežino	šeimai, giminei	neturi nuomonės	ne	1
vyras	62	aukštesnysis	taip	taip	visiškai sutinka	kažką žino	tik šeimai	teigiamai	galbūt	1
moteris	43	aukštesnysis	taip	taip	visiškai sutinka	kažką žino	šeimai, giminei	teigiamai	nesusimąsto	1
moteris	65	aukštas	taip	taip	visiškai sutinka	žino	neturi nuomonės	teigiamai	ne	1
vyras	32	aukštas	taip	ne	nei sutinka, nei nesutinka	kažką žino	tik šeimai	neturi nuomonės	nesusimąsto	1
moteris	31	aukštas	taip	taip	labiau sutinka, nei nesutinka	žino	tik šeimai	teigiamai	nesusimąsto	2
vyras	20	vidurinis	taip	taip	visiškai sutinka	žino	tik šeimai	teigiamai	taip	1
moteris	19	vidurinis	taip	taip	visiškai sutinka	kažką žino	šeimai, giminei, draugams	teigiamai	nesusimąsto	1

1	2	3	4	5	6	7	8	9	10	11
vyras	29	aukštasis	taip	taip	visiškai sutinka	kažką žino	tik šeimai	teigiamai	ne	1
vyras	25	n. aukštasis	taip	taip	visiškai sutinka	kažką žino	šeimai, giminei	teigiamai	galbūt	1
moteris	24	aukštasis	taip	taip	labiau sutinka, nei nesutinka	kažką žino	tik šeimai	teigiamai	nesusimąsto	3
vyras	26	aukštesnysis	taip	taip	visiškai sutinka	kažką žino	tik šeimai	teigiamai	galbūt	2
vyras	23	n. aukštasis	taip	taip	visiškai sutinka	kažką žino	šeimai, giminei	teigiamai	galbūt	1
vyras	40	aukštasis	taip	taip	visiškai sutinka	žino	tik šeimai	teigiamai	galbūt	2
vyras	50	aukštasis	taip	taip	nei sutinka, nei nesutinka	kažką žino	šeimai, giminei, draugams	teigiamai	ne	7
vyras	37	vidurinis	taip	ne	labiau sutinka, nei nesutinka	kažką žino	tik šeimai	neturi nuomonės	nesusimąsto	1
moteris	24	aukštesnysis	taip	taip	visiškai sutinka	žino	tik šeimai	teigiamai	galbūt	1
vyras	21	vidurinis	taip	taip	visiškai sutinka	žino	šeimai, giminei, draugams	teigiamai	galbūt	4
moteris	20	profesinis	taip	taip	visiškai sutinka	žino	tik šeimai	neturi nuomonės	ne	2
moteris	44	aukštasis	taip	neturi n	visiškai sutinka	kažką žino	šeimai, giminei, draugams	teigiamai	nesusimąsto	1

1	2	3	4	5	6	7	8	9	10	11
vyras	47	aukštasis	taip	taip	visiškai sutinka	žino	tik šeimai	teigiamai	nesusimąsto	1
vyras	52	aukštasis	taip	taip	visiškai sutinka	kažką žino	tik šeimai	neturi nuomonės	nesusimąsto	1
moteris	48	vidurinis	taip	taip	labiau sutinka, nei nesutinka	kažką žino	be abejo	teigiamai	nesusimąsto	1
moteris	58	aukštesnysis	taip	ne	nei sutinka, nei nesutinka	kažką žino	neaukotų	neturi nuomonės	ne	6
vyras	31	vidurinis	taip	taip	visiškai sutinka	kažką žino	neturi nuomonės	neturi nuomonės	nesusimąsto	4
vyras	29	aukštasis	taip	taip	visiškai sutinka	žino	be abejo	teigiamai	taip	1
moteris	51	aukštesnysis	taip	taip	labiau sutinka, nei nesutinka	kažką žino	tik šeimai	teigiamai	ne	1
vyras	38	n. aukštasis	taip	taip	labiau sutinka, nei nesutinka	kažką žino	tik šeimai	teigiamai	ne	1
moteris	22	n. aukštasis	taip	taip	visiškai sutinka	žino	tik šeimai	teigiamai	galbūt	1

Pradinė kintamųjų analizė

Pagrindinė pradinės duomenų analizės idėja yra ištirti požymius ir nustatyti, kurios stebimų požymių savybės pasitaiko dažnai, o kurios rečiau arba iš viso nepasitaiko. Kitaip sakant, reikia atskirti būdingas požymių savybes nuo nebūdingų, netipiškų.

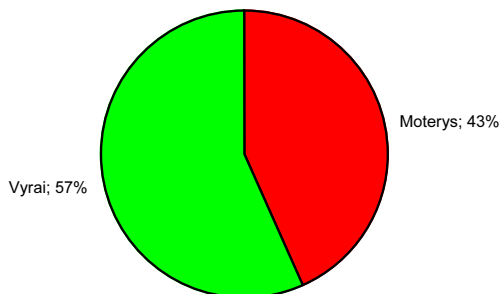
Lytis. Pirmasis tiriamas požymis yra lytis. Šis požymis yra matuojamas pavadinimų arba nominaliojoje skalėje. Požymį „lytis“ sudaro tik dvi galimos reikšmės – „vyras“ arba „moteris“. Išsiaiškinama, kurią stebimų dalį sudaro vyrai, o kurią – moterys, t. y. šių reikšmių dažniai. Gauti rezultatai rašomi į dažnių lentelę (2.5 lentelė). Šios lentelės pirmajame stulpelyje pateikiamas požymio pavadinimas (Lytis) ir jo įgyjamos reikšmės (Moteris, vyras). Antrajame stulpelyje „Dažnis“ rašomi kiekvienos požymio reikšmės dažniai. Taigi iš 30 apklaustų asmenų 13 buvo moteriškos lyties, 17 – vyriškos. Trečiajame stulpelyje rašomos sukaupusių dažnių reikšmės. Sukaupieji dažniai – tai dažnių sumos. Ketvirtajame stulpelyje pateikiamos dažnių procentinės reikšmės, paskutiniajame stulpelyje – procentinės sukaupusių dažnių reikšmės.

2.5 lentelė. Tiriamų asmenų pasiskirstymas pagal lytį

Lytis	Dažnis (Count)	Sukauptasis dažnis (Cumul. Count)	Dažnis procentais (Percent)	Sukauptasis dažnis procentais (Cumul. Percent)
Moteris	13	13	43,33	43,33
Vyras	17	30	56,67	100,00
Missing	0	30	0	100

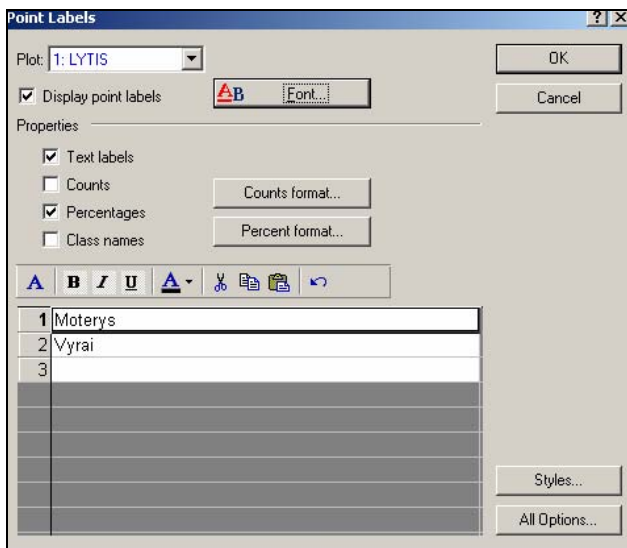
2.5 lentelės duomenis pateiksime grafiškai. Duomenys, išmatuoti nominaliaja skale, dažniausiai vaizduojami stulpelinėmis (*bar charts*, *columns charts*) ir skritulinėmis (*pie charts*) diagramomis. Labiausiai mėgstamas nominaliųjų duomenų vaizdavimo būdas – skritulinės diagramos (2.7 pav. a). Šio tipo diagramos patogios tada, kai požymio reikšmių skaičius yra nedidelis. Vaizduojant duomenis skritulinio tipo diagramomis galima pasirinkti, kaip diagramoje bus pateikiamos požymio reikšmės: šalia diagramos rašomi dažniai procentais (*Percentages*), dažnių reikšmių skaičius (*Count*), reikšmių žymėjimas (*Text labels*), kategorijų pavadinimai (*Class names*) (2.7 pav. b).

a)



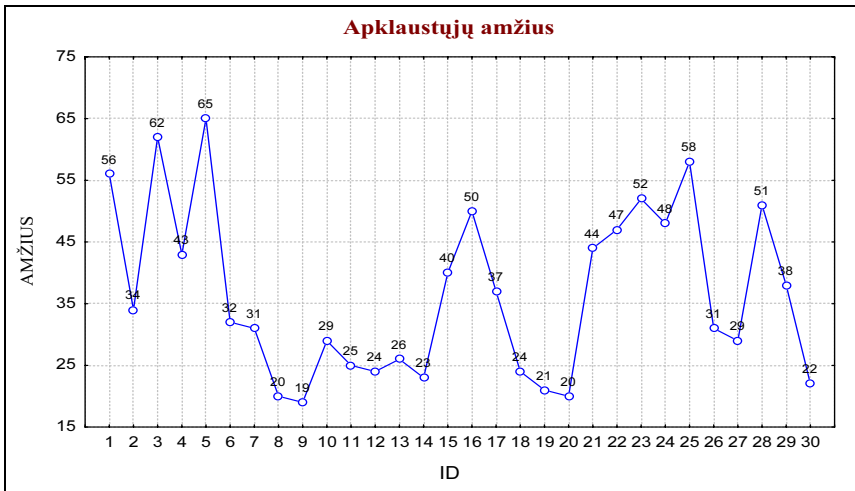
LYTIS

b)



2.7 pav. Vyrų ir moterų proporcijos skritulinėje diagramoje (a) ir parametrų parinkimo langas (b)

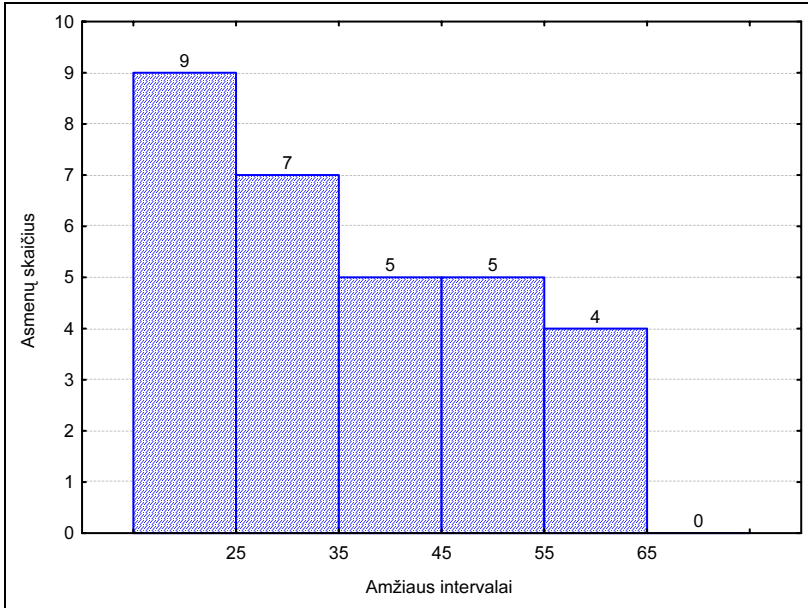
Gimimo metai arba amžius. Antrasis tiriamas požymis yra gimimo metai. Atliekant apklausas, rekomenduojama klausti ne žmogaus amžiaus, o jo gimimo metų. Iš gimimo metų apskaičiuojamas tiriamų asmenų amžius. Amžius matuojamas kiekybine (skaitine) skale. Pradedant analizuoti kiekybinius stebėjimus, pirmiausia nubraižoma linijinio tipo (*linear*) diagrama, kurioje geriausiai matyti padarytos klaidos, pavyzdžiui: jei užrašant duomenis netyčia vietoje 30 metų parašoma 300 metų.



2.8 pav. Apklaustųjų amžiaus linijinė diagrama

Pateikta linijinė diagrama (2.8 pav.) nėra labai informatyvi – tai tik bendras išmatuotų reikšmių pavidalas. Žiūrint į šią diagramą sunku pasakyti, kokio amžiaus asmenų buvo daugiausia, o kokio – tik atskiri atvejai. *Linijinės diagramos* dažniau braižomos tada, kai reikia iliustruoti požymio kitimą *laike*.

Kitas, informatyvesnis, kiekybinių stebėjimų vaizdavimo būdas – grupinių duomenų stulpelinė diagrama – histograma. Kai kiekybinių stebėjimų yra daug, duomenis tenka grupuoti. Grupuojant parenkamas intervalo ilgis, ir visi duomenys, kurie patenka į vieno intervalo ribas, yra vienos grupės (2.9 pav.).

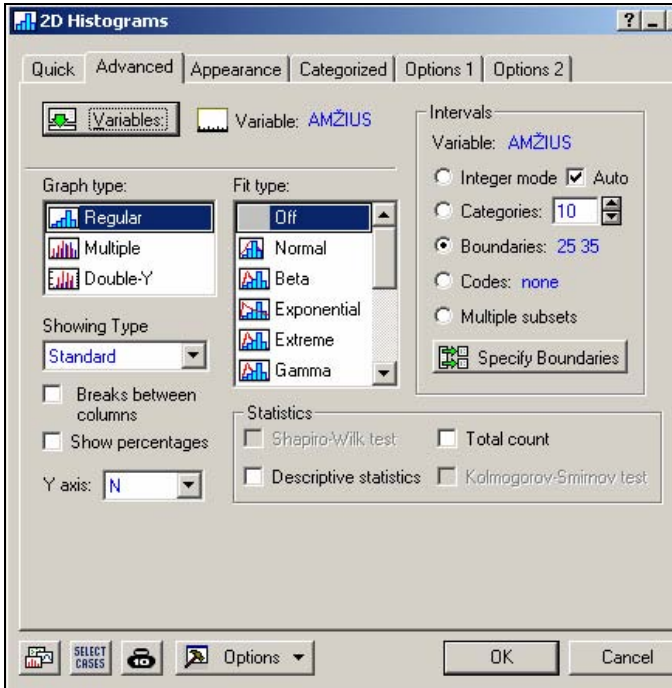


2.9 pav. Respondentų amžiaus grupinių duomenų stulpelinė diagrama – histograma.

2.9 pav. parinktas intervalo ilgis – 10 metų, tačiau pirmojo intervalo ilgis yra apibrėžtas tik iš vienos pusės – jame pavaizduota, kiek yra asmenų, kurių amžius neviršija 25 metų. Intervalo ilgis nustatomas pagal bendras taisykles. Beveik į kiekvieną intervalą turi patekti vidutiniškai ne mažiau kaip penki duomenys, o intervalų turėtų būti ne mažiau kaip 5 ir ne daugiau kaip 15. Pageidautina, kad intervalai būtų vienodo ilgio, kad gautas grupės būtų galima palyginti. Skaičiuojant statistinėmis programomis, yra daug galimybių intervalų skaičiui arba intervalo ilgiui parinkti (2.10 pav.).

Dešinėje lango pusėje (*Intervals*) galima nurodyti būdą, kaip suskirstyti duomenis į intervalus. Pirmuoju siūlomu būdu (*Integer Mode*) nustatoma, kad intervalų ribos yra sveikieji skaičiai, o intervalų skaičių parenka pati programa. Antruoju būdu nurodoma pageidautinas intervalų skaičius, o ribas apskaičiuoja statistinė programa. Naudojant trečiąjį būdą, vartotojas nurodo tiksliai viršutinių intervalų ribų reikšmes. Šiuo

būdu gauta diagrama pavaizduota 2.9 pav. Ketvirtuoju būdu galima atrinkti norimas reikšmes, nurodant jų kodus. Paskutiniuoju būdu, naudojant loginę *IF* funkciją, galima sudaryti logines išraiškas, pagal kurias bus atrinkti duomenys diagramai.



2.10 pav. *Statistica* programos intervalų savybių parinkimo langas

Jeigu duomenų daug, grupių skaičiui n nustatyti kartais pritaikoma formulė:

$$k=1+3,322 \log(n);$$

čia: k – intervalų skaičius, n – visų duomenų skaičius, \log – dešimtainis logaritmas.

Požymio „Amžius“ padėties ir sklaidos charakteristikos

Kadangi požymis „Amžius“ yra išmatuotas *kiekybine* skale, apskaičiuojamos jo pagrindinės charakteristikos (2.6 lentelė).

2.6 lentelė. Požymio „Amžius“ charakteristikų rezultatai

	Imties dydis (Valid N)	Vidurkis (Mean)	Mediana (Median)	Minimali reikšmė (Minimum)	Maksimali reikšmė (Maximum)	Dispersija (Variance)	Standartinis nuokrypis (Std.Dev.)
AMŽIUS	30	36,7	33	19	65	190,7	13,809

Iš 2.6 lentelės matyti, kad apklaustųjų amžiaus vidurkis yra 36,7 metų, o mediana – 33 metai. Jauniausias apklaustas asmuo buvo 19 metų, o vyriausias – 65 metų. Standartinis nuokrypis, kuris rodo duomenų sklaidą apie vidurkį, apytiksliai yra 14 metų; tai rodo, kad vidutiniškai kiekvieno asmens amžius yra nutolęs nuo vidurkio per 14 metų.

Išsilavinimas. Trečias požymis yra išsilavinimas. Šis požymis yra tipiškas rangų skalės kintamasis. Atsakant į klausimą apie respondento išsilavinimą, anketoje pateikti atsakymų tipai: nebaigtas vidurinis, vidurinis, profesinis, aukštesnysis, nebaigtas aukštesnysis, aukštasis, nebaigtas aukštasis. Kaip šio požymio reikšmės buvo pasiskirsčiusios tarp respondentų? Šiuo tikslu sudarysime dažnių lentelę (2.7 lentelė).

2.7 lentelė. Tiriamų asmenų pasiskirstymas pagal išsilavinimą

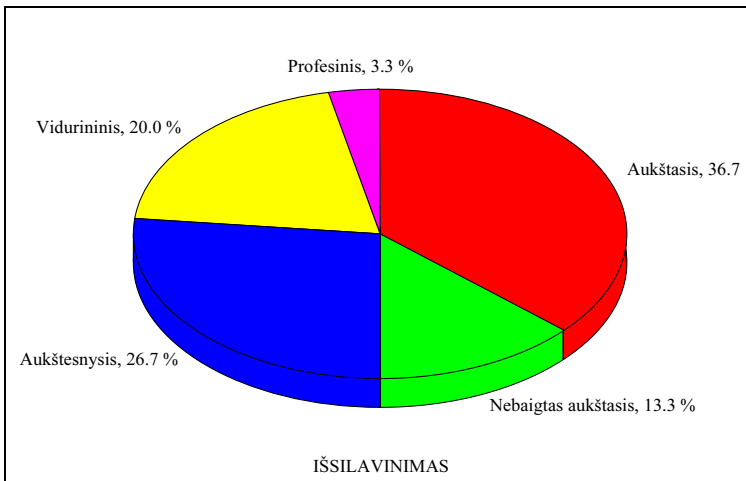
Išsilavinimas	Dažnis	Sukauptasis dažnis	Dažnis procentais	Sukauptasis dažnis procentais
	(Count)	(Cumul. Count)	(Percent)	(Cumul. Percent)
Aukštasis	11	11	36,667	36,667
Nebaigtas aukštasis	4	15	13,333	50
Aukštesnysis	8	23	26,667	76,667
Vidurinis	6	29	20	96,667
Profesinis	1	30	3,333	100
Trūksta reikšmių (Missing)	0	30	0	100

Iš 2.7 lentelėje pateiktų dažnių matyti, jog tarp apklaustųjų nebuvo asmenų su nebaigtu viduriniu bei nebaigtu aukštesniojo išsilavinimu.

Kitų apklaustųjų išsilavinimas pagal proporciją yra toks: su aukštuoju išsilavinimu 11, arba 36,7 proc. asmenų; su nebaigtu aukštuoju – 4, arba 13,3 proc.; su aukštesnioju – 8, arba 26,7 proc.; su viduriniu – 6, arba 20 proc.; su profesiniu – 1, arba 3,3 proc. Paskutinėje 2.7 lentelės eilutėje pateikta, kiek respondentų į šį klausimą neatsakė. Šiame tyrime tokių asmenų nebuvo, todėl ir dažnis yra nulis.

Daugiausia tyrime dalyvavo asmenų su aukštuoju bei aukštesnioju išsilavinimu. Vidurinį išsilavinimą turinčių žmonių taip pat daug, tačiau reikia atsižvelgti į tai, jog dauguma jų dabar studijuoja aukštosiose mokyklose.

Šį požymį grafiškai galima pavaizduoti skrituline arba sugrupuotų duomenų stulpeline diagrama – histograma. Kadangi asmenų su nebaigtu viduriniu ar nebaigtu aukštesnioju išsilavinimu nebuvo, reikšmių sąrašas sumažėjo iki penkių reikšmių. Statistinės programos siūlo keturis skritulinių diagramų tipus – dvimatę skritulinę, dvimatę elipsinę, trimatę skritulinę ir trimatę elipsinę. 2.11 pav. išsilavinimas pavaizduotas trimatėje elipsinė diagrama. Vyrų ir moterų dažniams pavaizduoti (2.7 pav.) buvo pasirinkta dvimatė skritulinė diagrama. Teikiant tyrimo rezultatus, galima rinktis bet kurį norimą diagramos tipą, tačiau nereiktų šių tipų kaitalioti (t. y. pasirinkus dvimatę skritulinę diagramą, ir kitos skritulinės diagramos turėtų būti to paties tipo).



2.11 pav. Išsilavinimas pavaizduotas trimatėje elipsinė diagrama

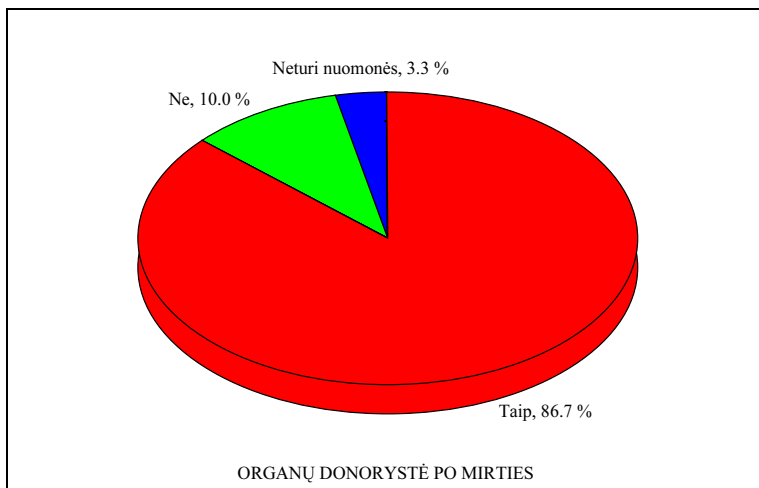
Į ketvirtą klausimą „*Ar žinote apie tokį reiškinį, kaip donorystė?*“ teigiamai atsakė visi apklaustieji respondentai, t. y. visi 30 apklaustųjų asmenų į klausimą, ar žino apie tokį reiškinį, kaip donorystė, pasirinko atsakymą „taip“. Todėl šio požymio nenagrinėsime.

Į penktą klausimą „*Ar sutiktumėte paaukoti organus po savo mirties?*“ teigiamai atsakė 86,7 proc. asmenų, 3,3 proc. asmenų neturėjo nuomonės, o 10 proc. asmenų teigė, jog to daryti nesutiktų (2.8 lentelė).

2.8 lentelė. Požiūris į organų paaukojimą po mirties

	Dažnis	Sukauptasis dažnis	Dažnis procentais	Sukauptasis dažnis procentais
	(Count)	(Cumul. Count)	(Percent)	(Cumul. Percent)
Taip	26	26	86,667	86,667
Ne	3	29	10	96,667
Neturi nuomonės	1	30	3,333	100
Trūksta reikšmių (Missing)	0	30	0	100

Iš 2.12 pav. matyti, kad dauguma apklaustųjų sutiktų paaukoti savo organus po mirties.



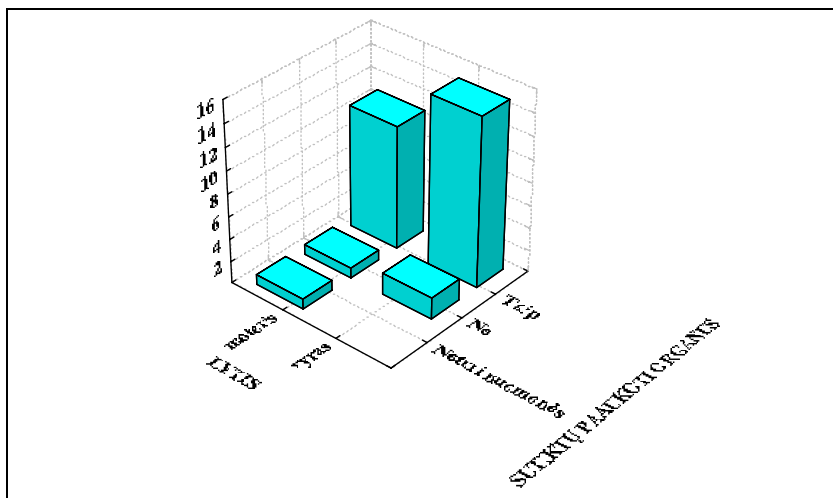
2.12 pav. Organų paaukojimas po mirties

Kadangi jau turime keletą kokybinių – nominaliųjų ir ranginių – kintamųjų, galima sudaryti porines dažnių lenteles. Apskaičiuokime, ar vyrų ir moterų, sutinkančių paaukoti savo organus po mirties, dažniai yra vienodi. Šiuo tikslu sudarome dviejų požymių „Lytis“ ir „Organų paaukojimas po mirties“ bendrąją dažnių lentelę (2.9 lentelė).

2.9 lentelė. Požymių „Lytis“ ir „Organų paaukojimas po mirties“ bendroji (porinė) dažnių lentelė

	Sutiktų	Nesutiktų	Neturi nuomonės	Totals
Moteris	11	1	1	13
Vyras	15	2	0	17
All Grps	26	3	1	30

Iš 2.9 lentelės matyti, kad iš 13 moterų sutiktų paaukoti organus 11 moterų, nesutiktų – 1, neturi nuomonės – 1. Iš 17 vyrų 15 sutiktų organus paaukoti, 2 – ne. Šie rezultatai pateikiami grafiškai (2.13 pav.).



2.13 pav. Požymių „Lytis“ ir „Organų paaukojimas po mirties“ diagrama

I šeštą klausimą „*Ar sutinkate su teiginiu, jog organų donorystė labai reikalinga?*“ buvo pateikti penki galimi atsakymai: visiškai sutinku; labiau sutinku nei prieštarauju; nei sutinku, nei nesutinku; labiau ne-

sutinku nei sutinku; visiškai nesutinku. Tai ranginio tipo kintamasis. Jo reikšmių dažniai pateikti 2.10 lentelėje.

2.10 lentelė. Požiūris į donorystės reikalingumą

	Dažnis	Sukauptasis dažnis	Dažnis procentais	Sukauptasis dažnis procentais
	(Count)	(Cumul. Count)	(Percent)	(Cumul. Percent)
Visiškai sutinka	20	20	66,7	66,7
Labiau sutinka nei prieštarauja	6	26	20,0	86,7
Nei sutinka, nei nesutinka	3	29	10,0	96,7
Visiškai nesutinka	1	30	3,3	100,0
Trūksta reikšmių (<i>Missing</i>)	0	30	0	100

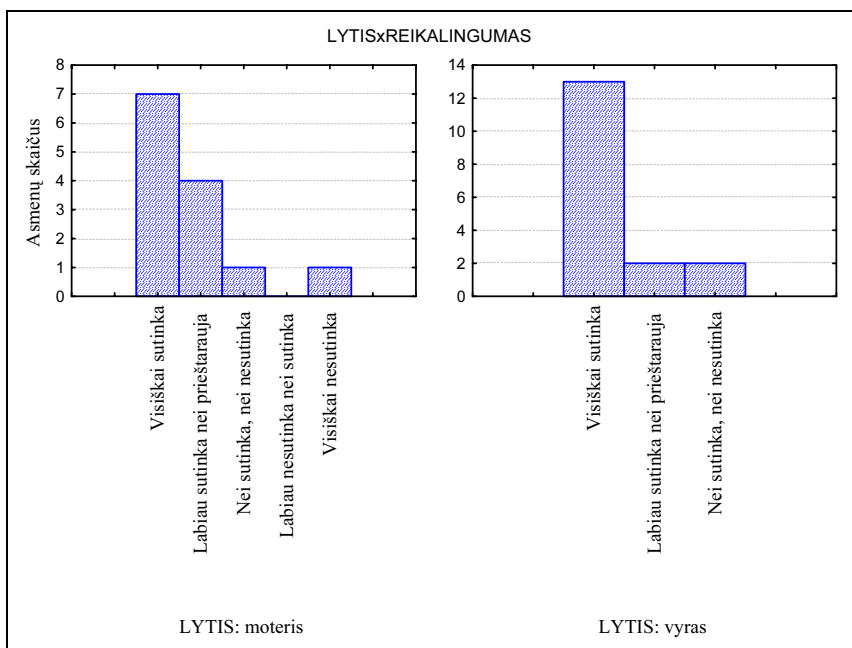
Iš 2.10 lentelės duomenų galima spręsti, jog 86,7 procento apklaustųjų mano, kad donorystė reikalinga (paskutinis stulpelis „Sukauptasis dažnis procentais“), 10 procentų neturi nuomonės, o 3,3 procento mano, kad organų donorystė yra nereikalinga. 2.10 lentelėje pateiktus dažnius pavaizduosime stulpeline diagrama (2.14 pav.).



2.14 pav. Nuomonių apie donorystės reikalingumą pasiskirstymas

Iš 2.13 pav. matyti, kad 20 apklaustųjų su šiuo teiginiu visiškai sutinka, 6 – labiau sutinka nei prieštarauja, 3 – neturi nuomonės, visiškai nesutinka tik 1. Iš šių duomenų galima daryti išvadą, jog daugiau apklaustųjų tam tikra prasme su šiuo teiginiu sutinka: 26 asmenys iš 30 pasirinko teigiamus atsakymus.

Išsiaiškinkime, ar nuomonė apie donorystės reikalingumą priklauso nuo lyties. Šiuo tikslu sudarykime porinę dažnių lentelę, o skaičiavimo rezultatus pateikime diagramomis. Kadangi šiuo atveju analizuojami du kintamieji, turintys po keletą reikšmių, juos geriau pavaizduoti ne trimatėje diagrama, o suskirstyti grupėmis (2.15 pav.)



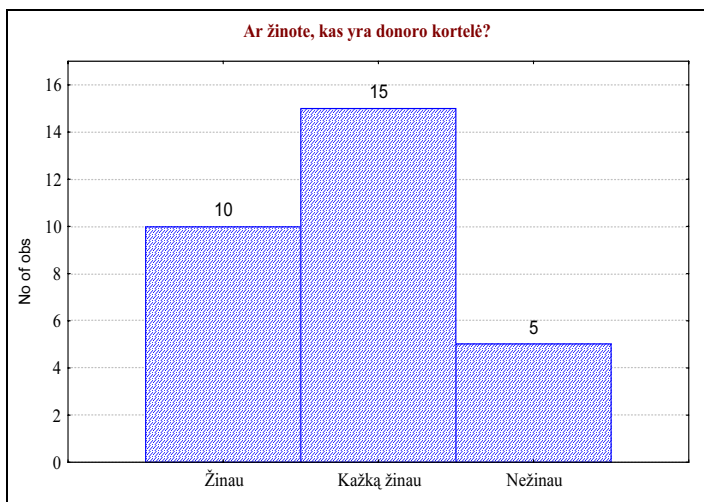
2.15 pav. Nuomonės apie donorystės reikalingumą pasiskirstymas pagal respondentų lytį

2.15 pav. diagrama aiškiai atskleidžia, jog abiejų lyčių grupėse vyrauja teigiama nuomonė apie organų donorystės reikalingumą.

I septintą klausimą „*Ar žinote kas yra donoro kortelė?*“ teigiamai atsakė 10, arba 33,33 proc. respondentų. Kažką yra girdėję, bet tiksliai nežino pusė respondentų. Taigi pusė atsakymų, t. y. 50,0 proc. buvo tei-

giami, respondentai apie donoro kortelės žino arba yra kažką girdėję. Nedaug respondentų – 5, arba 16,67 proc., visai nėra girdėję apie donoro kortelę. Kadangi į šį klausimą buvo pateikti tik trys galimi atsakymai: žinau; nežinau; kažką esu girdėjęs, bet tiksliai nežinau, tai dažnių lentelė ir histograma yra nedidelės, ir jas galima pateikti kartu (2.16 pav.)

	Dažnis	Sukauptasis dažnis	Dažnis procentais	Sukauptasis dažnis procentais
Žinau	10	10	33,333	33,333
Kažką žinau	15	25	50	83,33
Ne, nežinau	5	30	16,667	100



2.16 pav. Donoro kortelės populiarumas

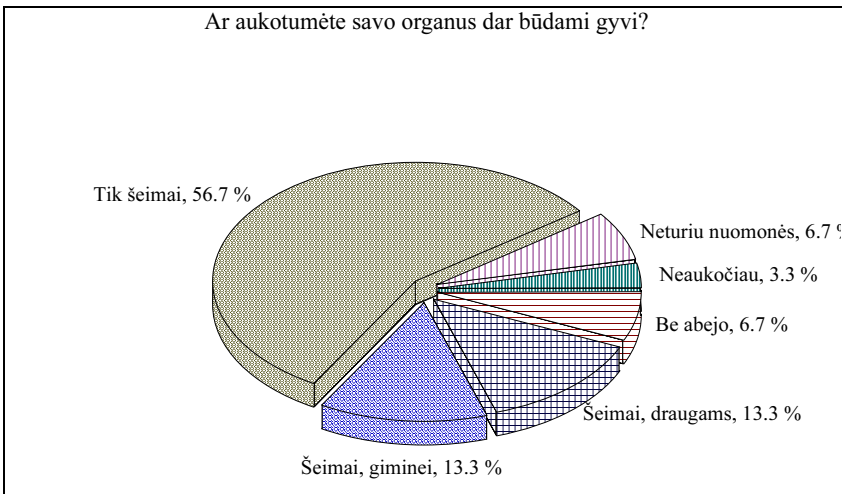
Pirmas, antras ir trečias klausimas yra personaliniai, ir iš jų sužinojome respondento lytį, amžių ir išsilavinimą. Klausimai nuo ketvirto iki aštunto buvo pateikiami norint sužinoti respondento nuomonę ir žinias. Atsakant į aštuntą klausimą, respondentui tenka rinktis gyvenimišką poziciją – aukoti ar neaukoti organus. Todėl į šį klausimą „*Ar aukotumėte kokį nors organą dar būdamas gyvas?*“ pateikiama daugiau ir įvairesnių atsakymų variantų: taip, be abejo; aukočiau šeimai, giminaičiams ir draugams; aukočiau šeimai bei giminaičiams; aukočiau tik artimiausiems šeimos nariams; tikrai neaukočiau; neturiu nuomonės. Tokie atsa-

kymai apklausiamiems žmonėms leidžia teikti prioritetus, o ne viena-reikšmišką nuomonę, pavyzdžiui, taip arba ne. Šio požymio dažniai pateikti 2.11 lentelėje.

2.11 lentelė. Pasiryžimas aukoti organą dar esant gyvam

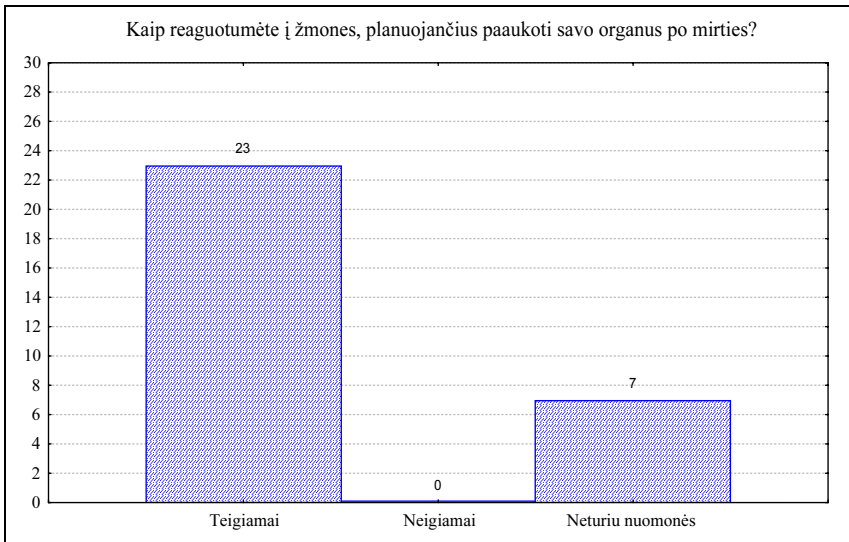
	Dažnis	Sukauptasis dažnis	Dažnis procentais	Sukauptasis dažnis procentais
Be abejo	2	6	6,667	6,667
Šeimai, giminei, artimiesiems draugams	4	6	13,333	20,000
Šeimai, giminei	4	10	13,333	33,333
Tik šeimai	17	27	56,667	90,000
Neaukotų	1	28	3,333	93,333
Neturi nuomonės	2	30	6,667	100,000
Trūksta (<i>Missing</i>)	0	30	0	100

Lygindami šiuos dažnius su pateiktais 2.8 lentelėje (apie organų aukojimą po mirties), galime pastebėti didelius pasirinktų atsakymų skirtumus ir tam tikrus nesutapimus. Savo organų po mirties nesutiktų paaukoti 10 proc. žmonių, o neaukotų dar būdami gyvi – 3,3 proc. Grafiškai šie rezultatai pateikti 2.17 pav.



2.17 pav. Aukotų savo organus būdamas gyvas

Devintuoju klausimu norima sužinoti respondento nuomonę apie organų donorystę po mirties: „**Kaip reaguojate\reaguotumėte į žmogų, planuojantį atiduoti savo organus po mirties?**“ Kadangi apie 87 procentų apklaustųjų teigiamai žiūri į organų donorystę po mirties (2.18 pav.), turėtume tikėtis panašių atsakymų ir į šį klausimą.

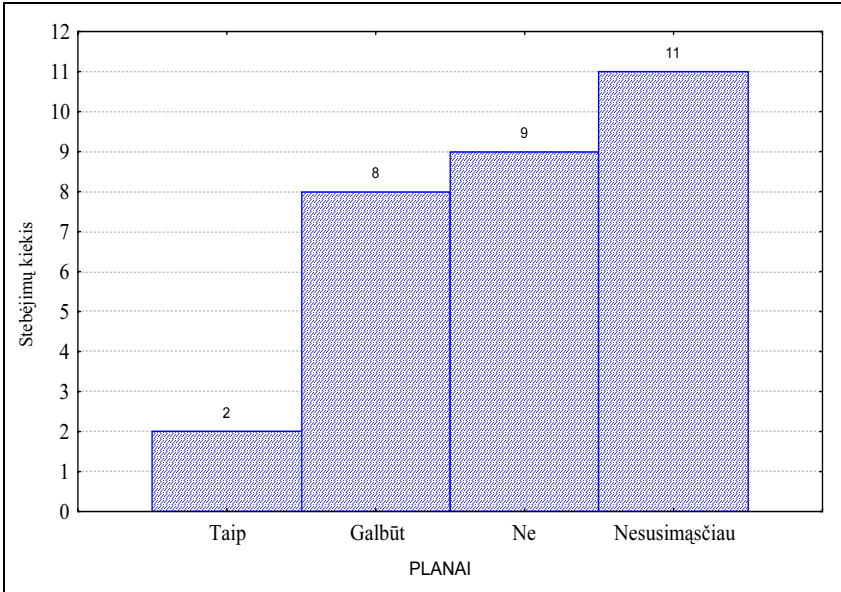


2.18 pav. Požiūris į organų donorystę po mirties

Į dešimtą klausimą „**Ar artimiausiu metu planuojate įsigyti donoro kortelę?**“ buvo pateikti keturi galimi atsakymai: taip, galbūt, ne, nesu-simąščiau apie tai. Šio ranginio tipo kintamojo reikšmių dažniai pateikti historigrameje 2.19 pav.

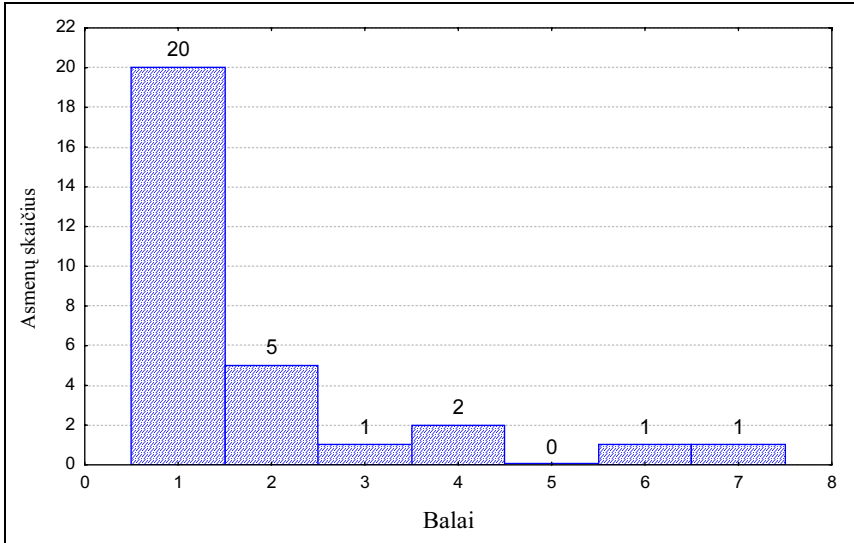
Nepaisant daugumos respondentų teigiamo požiūrio į organų dono-rystę, tik du asmenys apsisprendę artimiausiu metu įsigyti donoro korte- lę (2.19 pav.). Įdomu tai, kad tik 2 (arba 6,67 proc.) iš apklaustųjų į šį klausimą atsakė teigiamai, nors ankstesniuose mūsų tyrimo rezultatuose matyti, jog net 2/3 apklaustųjų į žmones, ketinančius tapti donorais, rea- guoja teigiamai, ir net 86,7 proc. yra pasiryžę savo organus paaukoti po mirties. Daugiausia respondentų, net 36,67 proc., teigė apie tai negalvo-

jantys ir pasirinko atsakymą „nesusimąščiau apie tai“. Nemažai respondentų to daryti neketina – 30,0 proc., o likusieji 26 pasirinko atsakymą „galbūt“.



2.19 pav. Planai apie ketinimą artimiausiu metu įsigyti donoro kortelę

Vienuoliktuoju klausimu apklaustųjų buvo prašoma dešimties balų skalėje pažymėti skaičių, geriausiai apibūdinantį žmonių, laukiančių organų persodinimo, padėtį: „Skalėje 1–10 įvertinkite žmonių, laukiančių transplantacijos, padėtį“. Skalė interpretuota taip: 1 – labai trūksta organų, 5 – organų pakanka, 10 – organų per daug. Tuo siekta išsiaiškinti, ar žmonės žino, kokios šiuo metu mūsų šalyje yra organų transplantacijos galimybės, ar žmonės mano, kad tinkamų persodinimui organų yra pakankamai. Net 20 asmenų organų transplantacijos laukiančių žmonių padėtį dešimties balų skalėje įvertino tik 1 balu (2.20 pav.), penki respondentai – 2 balais; vienas respondentai – 3 balais; du respondentai – 4 balais; penkiais balais – nė vienas; vienas – 6 balais; vienas – 7 balais.



2.20 pav. Žmonių, laukiančių transplantacijos, padėties vertinimas

Nors šis kintamasis yra ranginis, tačiau jo galimų įgyti reikšmių skaičius pakankamai didelis, kad būtų galima apskaičiuoti skaitines charakteristikas (2.12 lentelė).

2.12 lentelė. Žmonių, laukiančių transplantacijos, padėties vertinimo skaitinės charakteristikos

	Imties dydis	Vidurkis	Mediana	Minimumas	Maksimumas	Standartinis nuokrypis
Vertinimas	30	1,8	1	1	7	1,54

2.12 lentelėje nurodyta, kad vertinimų vidurkis yra 1,8 balo, o mediana – tik 1 balas, standartinis nuokrypis – 1,54. Vadinasi, gauti rezultatai yra „išsibarstę“.

Apklausoje rezultatai parodė, kad nors dauguma gyventojų pritaria organų donorystei, artimiausiu metu iš apklaustųjų asmenų tik du pla-

nuoja įsigyti donoro kortelę. Kalbant apie donoro korteles, galima teigti, jog Sveikatos apsaugos ministerijai ir kitoms atsakingoms institucijoms, reikėtų susirūpinti reklama bei kitaip skatinti žmones, jeigu norima padidinti organų donorų skaičių šalyje.

Užduotys

1.1. Lentelėje pateikti moksleivių atsakymai į anketos klausimus:

1. Lytis: 1 – vyras, 2 – moteris;
2. Amžius (metais);
3. Aš patenkintas(-a) savo gyvenimu: 1 – tikrai sutinku; 2 – sutinku; 3 – nesutinku; 4 – tikrai nesutinku.

Atlikite pradinę šių duomenų analizę:

- a) apskaičiuokite tinkamas duomenų charakteristikas;
- b) pavaizduokite duomenis grafiškai;
- c) paaiškinkite, kurie požymiai kiekybiniai, ranginiai, nominalieji. Kodėl?

ID	1. Lytis	2. Amžius	3. Gyvenimas	ID	1. Lytis	2. Amžius	3. Gyvenimas
1	Vyras	29	Sutinku	16	Moteris	22	Nesutinku
2	Vyras	26	Sutinku	17	Moteris	34	Sutinku
3	Vyras	30	Tikrai nesutinku	18	Vyras	26	Tikrai sutinku
4	Moteris	33	Sutinku	19	Vyras	28	Tikrai sutinku
5	Moteris	33	Sutinku	20	Vyras	32	Tikrai sutinku
6	Vyras	35	Sutinku	21	Vyras	31	Sutinku
7	Vyras	23	Nesutinku	22	Moteris	32	Sutinku
8	Vyras	29	Sutinku	23	Moteris	31	Tikrai nesutinku
9	Moteris	33	Sutinku	24	Vyras	35	T. sutinku
10	Moteris	26	Nesutinku	25	Vyras	26	Sutinku
11	Moteris	27	Nesutinku	26	Moteris	32	Nesutinku
12	Moteris	24	Sutinku	27	Moteris	30	Tikrai sutinku
13	Moteris	24	Sutinku	28	Vyras	57	Sutinku
14	Vyras	27	Tikrai sutinku	29	Vyras	39	Tikrai sutinku
15	Vyras	27	Nesutinku	30	Moteris	30	Sutinku

1.2. Pasirinkta tema sudarykite sociologinės apklausos anketą, kurioje būtų ne mažiau kaip po du kiekybinius, ranginius ir nominaliuosius požymius. Apklauskite atrinktus studentus ir surašykite duomenis. Pagal duomenų tipus parinkite tinkamas procedūras:

- a) apskaičiuokite: aritmetinį vidurkį, medianą, dispersiją, standartinį nuokrypį;
- b) sugrupuokite duomenis dažnių lentelėje;
- c) nubraižykite histogramas, linijines, stulpelines ar skritulines diagramas;
- d) atlikite porinę dažnių analizę ir rezultatus pavaizduokite diagramomis.



3. DEMOGRAFINĖ STATISTIKA

„Žmonės mažiausiai pastebi akivaizdžiausius dalykus.“

Aristotelis

Demografija yra mokslas, tiriantis gyventojų skaičiaus ir struktūros kitimą. Visų lygių valdymo procesuose vienas iš svarbiausių uždavinių yra įvertinti artimesnius ir tolimesnius gyventojų skaičiaus ir sudėties pokyčius. Tik įvertinus būsimą gyventojų skaičių ir sudėtį galima planuoti valstybės biudžeto ir savivaldybių pajamas, transporto infrastruktūrą, mokyklų ir ikimokyklinių įstaigų tinklo pokyčius, sudaryti efektyvias sveikatos ir šeimos planavimo programas.

Skirtingoms visuomenės grupėms būdinga skirtingas gyvenimo būdas ir skirtingi poreikiai. Ikimokyklinių įstaigų paslaugomis naudojasi vienos gyventojų grupės, mokyklomis – kitos, gerontologijos centrais – trečios ir t. t. Skirtingų gyventojų grupių ne tik skirtingas gyvenimo būdas, bet ir visų gyventojų bendri poreikiai – maistas, drabužiai, būstas. Pavyzdžiui: tipiniai maisto produktų krepšeliai skiriasi priklausomai nuo lyties, amžiaus, sveikatos būklės, atliekamo darbo pobūdžio.

Norint visuomenėje vykstančius procesus reguliuoti ir jų raidą padaryti optimalią, būtina suprasti, kad skirtingų gyventojų grupių ne tik skirtingi poreikiai, bet ir tai, kad visuomenėje vyksta nuolatinė konkurencinė kova dėl vartojimo resursų tiek socialinių grupių viduje, tiek tarp grupių. Gyventojų grupės, kurioms būdinga žemas konkurencingumo lygis, gali patekti į tokią padėtį, kai joms tenka kovoti už išlikimą. Tai sukelia agresiją, kuri gali tapti socialiniais neramumais. Todėl valstybės socialinė politika turi būti organizuota taip, kad šioms gyventojų grupėms (bedarbiams, vienišoms motinoms, daugiavaikėms šeimoms, pensininkams) būtų teikiama socialinė parama. Socialinė įtampa gali pasireikšti ne tik tarp gyventojų grupių, bet ir tarp gyventojų bei valdžios institucijų, jei pastarosios nesugeba tinkamai atlikti savo funkcijų.

Prie pagrindinių gyventojų skaičiaus ir struktūros požymių, kurie yra svarbūs analizuojant ir prognozuojant socialinius procesus, priski-

riami šie: išsilavinimas, profesija, užimamos pareigos, turimos pajamos ir nuosavybė bei šeimos pajamos. Socialinį asmens portretą atskleidžia demografiniai požymiai – lytis, amžius, šeimtinė padėtis, šeimos sudėtis ir etniniai požymiai – tautybė, religija, vartojamoji kalba.

3.1. Demografinių duomenų šaltiniai

Demografiniai duomenys gaunami dviem pagrindiniais būdais: gyventojų surašymu ir valstybės institucijų registruojama statistika (gimimai, mirtys, ištuokos, gyventojų migracija ir pan.). Jeigu valstybinės statistinės institucijos visiškai netenkina demografinės informacijos poreikio, tai informacijos trūkumą kompensuoja kiti šaltiniai: sociologijos institutai, visuomenės nuomonės tyrimo centrai ir pan.

Gyventojų surašymas atliekamas vienu metu visoje teritorijoje. Registruojamų rodiklių skaičius nėra pastovus. Pavyzdžiui: 1979 m. gyventojų surašyme surašymo lape buvo 16 klausimų, o 1989 m. – jau 25 klausimai.

Pirmuosius duomenis apie Lietuvos gyventojus buvo pradėta rinkti XIII amžiuje. Duomenys buvo reikalingi nustatyti, kiek yra tinkamų kariauti vyrų ir kiek žmonių gali mokėti mokesčius.

1790 m. surašytos visos tuometinės gyventojų grupės, todėl šį gyventojų surašymą galima laikyti pirmuoju visuotiniu Lietuvos Didžiojoje Kunigaikštystės gyventojų surašymu. Surašyti gyventojai buvo suskirstyti pagal lytį ir tikyba, o krikščionys papildomai suskirstyti pagal amžių ir pagal luomą (bajorai, dvasininkai, miestiečiai, valstiečiai). Taip pat 1790 m. surašymo metu buvo renkami duomenys apie natūralųjį gyventojų judėjimą: santuokas, gimimus ir mirimus.

1918 m. buvo paskelbta Lietuvos nepriklausomybė, o visuotinis gyventojų surašymas vyko 1923 m., nustačius valstybės sienas. Surašymas vyko 1923 m. rugsėjo 17–23 dienomis. Surašymo medžiagos apdorojimas (pagal iš anksto parengtą programą) truko visus 1924 metus ir keletą 1925 m. mėnesių. Rezultatų lentelės buvo publikuotos 1924 m. ir 1925 m. „Statistikos biuletenyje“.

Po Antrojo pasaulinio karo gyventojų surašymas įvyko 1959, 1970, 1979, 1989 m. (3.1 lentelė).

Atkūrus Lietuvos valstybės nepriklausomybę, 2001 m. įvyko visuotinis gyventojų ir būstų surašymas. 2001 m. spalio 30 dieną Statistikos

departamentas paskelbė pirmuosius duomenis: kiek surašyta gyventojų, kiek iš jų vyrų ir moterų, kaip jie išsidėstę Lietuvos teritorijoje (3.1 lentelė, paskutinė eilutė).

3.1 lentelė. Kai kurių Lietuvos gyventojų surašymų duomenys, nuolatiniai gyventojai (tūkst. gyv.)

Metai	Teritorija (tūkst. km ²)	Gyventojų skaičius (tūkst.gyv.)	Gyventojų tankumas 1 km ²
1897	91,2	2 536,0	38,9
1923	53,2	2 029,0	38,2
1959	65,2	2 696,7	41,6
1970	65,2	3 118,9	48,0
1979	65,2	3 391,5	52,1
1989	65,2	3 674,8	56,6
2001	65,3	3484,0	53,4

Šalia gyventojų surašymo ir valstybės institucijų registruojamos statistikos, atsižvelgiant į EUROSTAT reikalavimus, nuo 1994 m. Lietuvos statistikos departamentas atlieka *gyventojų užimtumo* tyrimus, o nuo 1996 m. – ir *namų ūkių biudžeto* tyrimus. Gyventojai apklausiami pagal specialias anketas, parengtas pagal EUROSTAT reikalavimus.

Darbo biržos skelbia informaciją apie bedarbius, įregistruotus darbo biržose, o Statistikos departamentas – apie nedirbančius asmenis, ieškančius darbo įvairiais būdais.

Namų ūkių biudžeto tyrimo objektas – namų ūkis, t. y. vienas asmuo arba asmenų grupė, kuri gyvena viename bute (name), turi bendrą biudžetą ir kartu maitinasi. Namų ūkiai atrenkami atsitiktinių imčių metodu iš gyventojų registro. Tokia atranka užtikrina vienodas visų visuomenės sluoksnių atstovų galimybes būti atrinktiems tyrimui. Namų ūkių biudžeto tyrime surenkama informacija apie namų ūkio narių sudėtį pagal amžių, lytį, išsimokslinimą, butų sąlygas, pajamas, išlaidas maisto ir ne maisto prekėms bei paslaugoms.

Namų ūkio biudžeto tyrimai tapo jungiamąja grandimi tarp atskirų tyrimų, kurie pateikdavo visuomenę tarsi tam tikrais „pjūviais“ (gyventojų skaičius, pagaminta produkcijos, suvartota produkcijos ir pan.), tačiau šių duomenų nebuvo galima gretinti ir palyginti. Taip ir likdavo ne-

aišku, kokios atskirų visuomenės sluoksnių pajamos, jų suvartojimas, gyvenimo sąlygos ir pan.

Kai kurią informaciją apie gyventojus galima gauti iš netiesioginių šaltinių, pavyzdžiui: moksleivių skaičiaus, studentų skaičiaus, muziejų lankytojų skaičiaus, įvairiomis transporto rūšimis pervežamų keleivių skaičiaus ir pan. Tačiau šie duomenys atspindi tik bendrą asmenų skaičių, be individualių charakteristikų.

Norint iš gausių demografinės informacijos šaltinių duomenų atrinkti naudingas ir reikalingas žinias, taikomi įvairūs demografiniai metodai. Iš šių metodų dažniausiai taikomi:

- pagrindinių vaisingumo, mirtingumo ir vedybinių charakteristikų skaičiavimas;
- aprašomieji demografinių duomenų metodai;
- amžiaus lentelių sudarymas;
- populiacijos projektavimas;
- duomenų integravimas ir procesų prognozavimas;

Tiriant gyventojų nuomonę, dažnai kartu atliekamas ir didesnės ar mažesnės apimties demografinis tyrimas. Pateikiame apklausos fragmentą, kurį panaudojo viešosios nuomonės ir rinkos tyrimų kompanija „TNS GALLUP“:

Anketa

DEMOGRAFINIAI KLAUSIMAI

D1. Vietovė, kurioje vyko apklausa:

RAJONAS _____

D2. Jūsų gyvenamosios vietovės dydis:

1. Per 200 tūkst. gyventojų
2. 30 – 200 tūkst. gyventojų
3. 2 – 30 tūkst. gyventojų
4. Iki 2000 gyventojų

D3. Respondento lytis:

1. Vyras
2. Moteris

D4. Kiek Jums metų?

_____ sukako metų

D5. Jūsų tautybė:

1. Lietuvis
2. Rusas
3. Lenkas
4. Kita

D6. Jūsų išsilavinimas:

1. Pradinis
2. Pagrindinis (10 klasių, anksčiau – 8 – 9)
3. Bendrasis vidurinis
4. Spec. vidurinis (PTM, technikumus)
5. Aukštasis, nebaigtas aukštasis

D7. Šiuo metu Jūs ... Tik vienas atsakymas

1. Pats sau darbdavys
2. Samdomas darbuotojas ⇒ *klauskite D8*
3. Dekretinėse, vaiko priežiūros atostogose ⇒ *D10*
4. Bedarbis, ieškote darbo ⇒ *klauskite D10*
5. Pensijoje ⇒ *klauskite D10*
6. Namų šeimininkė ⇒ *klauskite D10*
7. Mokinys, studentas ⇒ *klauskite D10*
8. Kita (*įrašykite*) _____

Klausimas D8 – samdomiems darbuotojams**D8. Jūsų pareigos:**

1. Aukščiausio lygio vadovas
2. Vidurinio lygio vadovas
3. Aukščiausio lygio specialistas (turintis aukštąjį išsilavinimą ir dirbantis pagal specialybę)
4. Vidurinio lygio specialistas, asistentas, techninis darbuotojas
5. Tarnautojas, prekybos bei paslaugų sferos darbuotojas, kariškis, apsaugos darbuotojas, policininkas ir pan.
6. Kvalifikuotas darbininkas (dirbantis su įranga ar mechanika; vairuotojas)
7. Darbininkas, dirbantis fizinį darbą

NB: Jei sunku nustatyti, kuo dirba respondentas, užrašykite jo pareigas išsamiai: _____

Klausimas D9 – visiems dirbantiems

D9. Jūs...

1. ...vedęs, ištekėjusi, gyvenate nesusituokę
2. ...viengungis, netekėjusi
3. ...gyvenate atskirai, išsiskyre, našlys, našlė

D10. Kiek žmonių, įskaitant ir Jus, gyvena kartu su Jumis ir tvarko bendrą ūkį? _____ žmonės

Jei respondentas gyvena vienas, toliau klauskite D14 klausimo. Kitiems užduokite D12 ir D13 klausimus.

D11. Kiek Jūsų šeimoje vaikų iki 16 metų amžiaus? _____ vaikai

D12. Ar yra Jūsų šeimoje vaikų ...?

1. 12 – 15 metų amžiaus
2. 7 – 11 metų amžiaus
3. 4 – 6 metų amžiaus
4. 2 – 3 metų amžiaus
5. 1 metų amžiaus
6. Jaunesnių negu 1 metų amžiaus

3.2. Demografiniai rodikliai ir grupės

Gyventojų populiacijos pokyčiai

Daugelis šalių, tiek išsivysčiusių, tiek ir besivystančių, patyrė didelius demografinius pokyčius. Reikšmingiausi pokyčiai būna tada, kai nuo didelio gyventojų prieaugio pereinama prie mažo. Iš pradžių šalyse mažėja mirtingumas ir didėja gimstamumas. Šie abu rodikliai kartu didina šalies gyventojų populiaciją. Po kurio laiko mirtingumo mažėjimas

sukelia gimstamumo mažėjimą. Šis žinomas reiškinys vadinamas demografiniu perėjimu (*demografic transition*).

Yra trys būdai (ir tik trys), kai populiacija gali keisti savo dydį ir struktūrą. Kiekvienas iš šių trijų būdų turėtų būti įvertintas kohortų komponentiniu modeliu. Šie trys būdai yra:

1. **Gimimai.** Natūraliausias gyventojų skaičiaus padidėjimo būdas – vaikų gimdymas. Šis rodiklis svarbus tuo, kad jis iliustruoja vieną iš pagrindinių demografinių principų. Tik tam tikra populiacijos dalis gali gimdyti vaikus (nepaisant visų medicinos mokslo pasiekimų). Kalbant demografiniais terminais, šio amžiaus moterų grupė pasižymi „gimdymo rizika“. Demografijoje „gimdymo rizika“ taikoma 10 – 49 metų moterims. Šio amžiaus moterims yra skaičiuojamas specialus *vaisingumo santykis* pagal amžiaus kohortas (grupes) 10–14, 15–19, 20–24, 25–29, 30–34, 34–39, 40–44, 44–49 metai.

2. **Mirtys.** Tai žmonių pasitraukimas iš populiacijos dėl mirties. Skirtingai nei gimdymas, visi žmonės pasižymi „mirties rizika“, todėl *mirtingumo santykis* yra skaičiuojamas visoms amžiaus/lyties kohortoms. Mirtingumo santykis yra palyginti aukštas pirmaisiais gyvenimo metais ir mažėja iki vidutinio amžiaus. Aukščiausias mirtingumo santykis yra senyvų žmonių amžiaus grupėse (65 m. ir vyresnių). Sudarant kohortų prognozės modelius, dažniausiai taikomas mirtingumo santykio padarinys – *išlikimo santykis*. Išlikimo santykis – tai tikimybė, kad asmuo, turintis tam tikrų savybių, išgyvens tam tikrą (konkretų) laiką.

3. **Migracija.** Dar vienas faktorius, turintis įtakos populiacijos dydžiui, yra migracija. Migracija gali vykti dviem kryptimis:

- **imigracija** – gyventojų atsikėlimas gyventi iš kitos šalies (regiono);
- **emigracija** – žmonių išvykimas gyventi iš vienos šalies (regiono) į kitą.

Ne taip, kaip gimimų ir mirčių, migracijos santykiai neskaičiuojami. Nustatyti, kas turi „migravimo riziką“ yra labai sunku, ypač imigrantams. Todėl vertinant migraciją dažniausiai vertinamas buvęs migracijos lygis, ir šiuo rodikliu naudojamosi projektuojant ateities imigracijos lygį. Į bet koki populiacijos didėjimą ar mažėjimą, nesusijusį su vaisingumo ir mirtingumo rodikliais, žiūrima kaip į migracijos rezultata.

Pavyzdžiui: jeigu per dešimtį metų užregistruota 1000 gimimų ir 800 mirčių, o populiacija padidėjo 400 asmenų, tai į šiuos papildomus 200 asmenų žiūrima kaip į migracijos rezultata.

Visus demografinius rodiklius galima suskirstyti į kelias pagrindines grupes: bendrieji demografiniai rodikliai, gimstamumo rodikliai ir mirtingumo rodikliai.

Bendrieji demografiniai rodikliai

Darbingo amžiaus gyventojai:

Lietuvoje iki 1995 m. – 16–60 m. vyrai ir 16–55 m. moterys; nuo 1995 m. vyrų darbingas amžius ilginamas 2 mėn. per metus ir moterų – 4 mėn. per metus, kol vyrai pasieks 65 m., o moterys – 60 metų.

Vaikais sveikatos statistikoje vadinami 0–14 metų gyventojai, o organizuojant sveikatos priežiūrą vaikams taikomas 15 metų kriterijus.

Gyventojų senėjimas apibūdinamas gyventojų senėjimo rodikliu (indeksu), kuris nustatomas keliais būdais. Jungtinių Tautų Organizacija rekomenduoja gyventojų senėjimo rodikliu laikyti 65 metų ir vyresnių žmonių skaičių (proc.) tarp visų gyventojų.

Gimstamumo rodiklis – gyvų gimusiųjų per metus skaičiaus santykis su vidutiniu metiniu gyventojų skaičiumi, padaugintas iš tūkstančio. Tai gimusių gyvų kūdikių 1000 gyventojų per metus rodiklis.

Mirtingumo rodiklis – mirusiųjų skaičius per metus 1000 gyventojų.

Natūralus gyventojų prieaugis – skirtumas tarp gimstamumo ir mirtingumo rodiklių 1000 gyventojų:

$$\frac{\text{Gyvų gimusiųjų skaičius} - \text{mirusiųjų skaičius per metus}}{\text{Vidutinis metinis gyventojų skaičius}} \times 1000$$

Santuokų (ištuokų) rodikliai – tai užregistruotų santuokų (ištuokų) skaičius per metus 1000 gyventojų.

Skaičiuojamas ir ištuokų skaičius 1000 santuokų.

Gimstamumo rodikliai

Vaisingumo rodiklis rodo, kiek gyvų kūdikių gimė 1000 vaisingo amžiaus (10–49 metų) moterų.

Vaisingumo rodikliai pagal motinos amžių – tai gyvų gimusiųjų skaičius 1000 atitinkamo amžiaus moterų.

Suminis vaisingumo rodiklis – tai vidutiniškai vienos moters pagimdytų vaikų per visą gyvenimą skaičius, jeigu visą tą laikotarpį būtų tie patys vaisingumo rodikliai, apskaičiuoti tiriamuoju laikotarpiu.

Bendrasis (bruto) reprodukcijos koeficientas parodo, kiek vidutiniškai mergaičių per visą gyvenimą pagimdytų viena moteris, jei visą tą laikotarpį išliktų vaisingumo rodikliai, skaičiuojami tiriamuoju laikotarpiu.

Grynasis (neto) reprodukcijos koeficientas parodo, kiek vidutiniškai viena moteris per visą gyvenimą pagimdytų mergaičių, kurios išgyvens iki to amžiaus, kokio būdamos motinos jas pagimdė.

Mirtingumo rodikliai

Mirtingumas dėl tam tikrų ligų (priežasčių) dažniausiai pateikiamas 100 000 gyventojų.

Mirties priežasčių struktūra – tai dėl tam tikros priežasties mirusiųjų skaičius (proc.) tarp visų priežasčių mirusiųjų skaičiaus.

Mirtingumas pagal žmonių amžių – tai tam tikros amžiaus grupės mirusių žmonių skaičiaus ir vidutinio metinio šio amžiaus žmonių skaičiaus santykis.

Tam tikro amžiaus gyventojų mirtingumas pagal mirties priežastis – tai tam tikro amžiaus dėl tam tikros priežasties mirusių žmonių skaičius (per metus), padalintas iš vidutinio metinio atitinkamo amžiaus žmonių skaičiaus ir padaugintas iš 100 000.

Kūdikių (vaikų iki vienerių metų) mirtingumas – tai per metus iš 1000 gimusių gyvų kūdikių mirusių kūdikių skaičius.

Vidutinė prognozuojama (numatoma) gyvenimo trukmė – tai vidutinis gyvenimo amžius (metai) esant skaičiuojamojo laikotarpio mirtingumo lygiui.

3.3. Struktūros dalių santykiniai dydžiai

Be absoliučių dydžių, dažnai skaičiuojamos sudarytų grupių dalys arba procentai, apibūdinantys atskirų visumos dalių santykinį „svorį“. Paprastai jie apskaičiuojami taikant procentinį santykį:

$$d = \frac{y_i}{\sum y} \cdot 100, \quad (3.1)$$

arba

$$d = \frac{\text{Visumos dalies dydis}}{\text{Suminis dydis}} \cdot 100 \cdot$$

Šalia santykinų „svorių“ dažnai apskaičiuojama, kiek kartų viena visumos dalis didesnė už kitą arba kiek vienos dalies elementų tenka 1, 10, 100, 1000 ar 100 000 kitos dalies elementų. Pavyzdžiui: vyrų ir moterų skaičiaus santykis gali būti pateiktas dviem būdais – gyventojų sudėtyje esančių vyrų ir moterų skaičius (proc.), arba moterų, tenkančių 1000 vyrų, skaičius. Skaičiai teigia, kad 2003 m. pradžioje Lietuvoje buvo 46,7 proc. vyrų ir 53,3 proc. moterų. Tūkstančiui vyrų teko 1141 moteris. Tačiau už šio bendro apibūdinimo slypi didelis vyrų ir moterų pasiskirstymo netolygumas pagal skirtingas amžiaus grupes.

Santykiniai intensyvumo dydžiai

Santykiniais intensyvumo dydžiais apibūdinamas vieno ar kito reiškinio išplitimas. Analizuojant demografinius gimimo, mirties, santuokų, ištuokų ir t. t. procesus, rodikliai skaičiuojami 1000 gyventojų ir pateikiami kaip metinės rodiklio reikšmės (per metus gimusių gyventojų skaičiaus ir pan.) santykis su vidutiniu metiniu gyventojų skaičiumi. Jei gaunamos reikšmės labai mažos, jos apskaičiuojamos 10 000 gyventojų.

Kadangi skaičiuojant intensyvumo rodiklius naudojamos vidutinės reikšmės, šie rodikliai kartais priskiriami prie centro (arba padėties) matų.

Sudėtingesnės santykinės charakteristikos

Demografinėje statistikoje taikomos ir sudėtingesnės apibendrinančios charakteristikos. Prie svarbių charakteristikų priskiriamas darbingo amžiaus gyventojų demografinis krūvis K_{dem} . Ši charakteristika apskaičiuojama pagal formulę:

$$K_{dem} = \frac{S_{0-14} + S_{60 \text{ ir daugiau}}}{S_{15-59}} \cdot 1000 \quad (3.2)$$

Čia: S_{0-14} – vaikų iki 14 m. skaičius; $S_{60 \text{ ir daugiau}}$ – gyventojų per 60 m. skaičius; S_{15-60} – darbingo amžiaus gyventojų skaičius.

Taikomos ir šios charakteristikos: K_v – vaikų skaičius, tenkantis 1000 darbingo amžiaus gyventojų ir K_p – pagyvenusių asmenų skaičius, tenkantis 1000 darbingo amžiaus gyventojų:

$$K_v = \frac{S_{0-14}}{S_{15-59}} \cdot 1000, \quad (3.3)$$

$$K_p = \frac{S_{60 \text{ ir daugiau}}}{S_{15-59}} \cdot 1000. \quad (3.4)$$

Charakteristikų K_v ir K_p suma lygi K_{dem} : $K_v + K_p = K_{dem}$.

Apskaičiavus šias charakteristikas pagal 2003 m. duomenis gauname, kad $K_{dem}=621$; $K_v=296$ ir $K_p=325$. Charakteristika K_{dem} rodo kad tūkstančiui 15–59 m. amžiaus gyventojų tenka 621 nedarbingo amžiaus asmenų. Apskaičiavus ankstesnių metų charakteristikas, galima stebėti jų kitimo tendencijas. Stebimi sisteminiai pokyčiai – mažėjantis vaikų skaičius, tenkantis 1000 darbingo amžiaus gyventojų ir didėjantis pagyvenusių asmenų skaičius, tenkantis 1000 darbingo amžiaus gyventojų. Tai visuomenės senėjimo požymiai. Toks visuomenės senėjimas veikia valstybės socialinę bei ekonominę padėtį. Valstybės socialinių garantijų dydis susijęs su tuo, kiek nedarbingo amžiaus asmenų tenka vienam darbingo amžiaus gyventojui.

3.4. Demografinių duomenų grafinis vaizdavimas

Įvairios diagramos, be abejonės, yra pats populiariausias duomenų vaizdavimo būdas. Informacija, pateikta bet kokio tipo lentelėje, visada gali būti pateikta grafiškai. Grafiniai vaizdai palengvina suvokti informaciją ir proceso kitimo tendencijas.

Vienas iš tradicinių demografinės statistikos grupavimo būdų – skirstymas į miesto ir kaimo gyventojus. Lietuvos statistikos departamentas pateikia šiuos duomenis (3.2 lentelė).

Miesto ir kaimo gyventojų santykis yra vienas iš greičiausiai kintančių demografinių rodiklių. Šis rodiklis jautriai reaguoja į socialinius

bei ekonominius ir politinius procesus. Miesto gyventojų daugiausia buvo 1991 m. (68,3 proc.). Miesto gyventojų mažėjimo tendencija atspindi valstybės ekonominę padėtį.

3.2 lentelė. Lietuvos miesto ir kaimo gyventojų skaičiaus kitimas

Vidutinis metinis gyventojų skaičius tūkstančiais				
Metai	Mieste	Kaime	Mieste (proc.)	Kaime (proc.)
1990	2520,3	1177,5	68,2	31,8
1991	2529,1	1175	68,3	31,7
1992	2520,9	1179,2	68,1	31,9
1993	2498,4	1184,2	67,8	32,2
1994	2472,3	1184,9	67,6	32,4
1995	2445,5	1183,6	67,4	32,6
1996	2430,8	1170,8	67,5	32,5
1997	2413,6	1161,6	67,5	32,5
1998	2387,8	1161,5	67,3	32,7
1999	2367,1	1157,1	67,2	32,8
2000	2345,6	1153,9	67,0	33,0
2001	2330,2	1151,1	66,9	33,1
2002	2321,7	1147,4	66,9	33,1

Grupiniai rodikliai dažnai pateikiami grafiškai. Tačiau ne visada lengva gauti informatyvų duomenų vaizdą. Jeigu 3.2 lentelės duomenis pavaizduosime tiesiogiai, diagramoje bus sunku išžiūrėti kitimo tendencijas. Norint aiškiau pavaizduoti nagrinėjamo proceso kitimą laike, dažnai taikomi *santykiniai* dydžiai.

Proceso kitimo santykinų dydžių tipai

Proceso kitimo santykinis dydis parodo proceso kitimą nagrinėjamoje laiko atkarpoje. Jis gaunamas to paties rodiklio (paprasto arba sudėtingo socialinio bei ekonominio rodiklio) reikšmę konkrečiu laikotarpiu ar momentu (y_1) palyginus su šio rodiklio reikšme ankstesniu laikotarpiu ar momentu (y_0). Jeigu lyginamos tik dvi reikšmės, gauta reikšmė yra kitimo koeficientas arba *indeksas*:

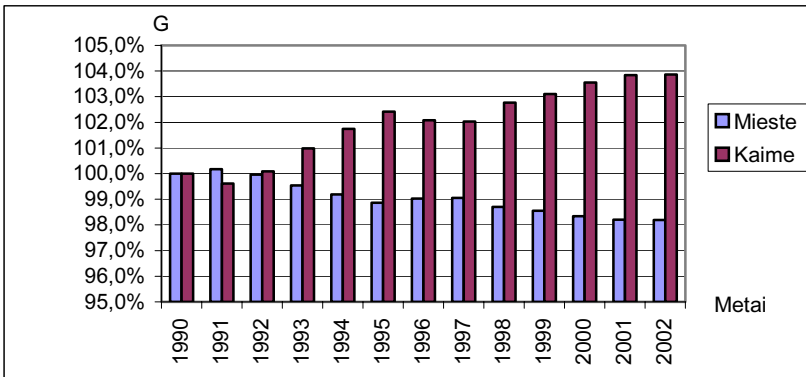
$$i = \frac{y_1}{y_0} . \quad (3.5)$$

Jeigu nagrinėjama keletas ar keliolika reikšmių, gaunamas dinaminis indeksas arba kitimo greitis. Kitimo greitis apskaičiuojamas dviem būdais: pirmuoju būdu nagrinėjamas lygis ($y_i, i=1,2,\dots,n$) palyginamas su pradiniu lygiu (y_0), ir nauja eilutė (santykinis kitimo greitis G) sudaroma taip:

$$G_i = \frac{y_i}{y_0} \cdot 100, \quad i = 1, 2, \dots, n. \quad (3.6)$$

Čia: santykinis kitimo greitis G_i išreikštas procentais.

Pagal 3.6 formulę apskaičiuotas vidutinis metinis miesto ir kaimo gyventojų skaičiaus santykinis kitimo greitis parodo, kaip kito miesto ir kaimo gyventojų skaičius pradinių metų (šiuo atveju 1990 m.) atžvilgiu. Diagramoje (3.1 pav.) abiejų rodiklių 1990 m. reikšmės prilyginamos arba vienetui, arba 100 procentų.



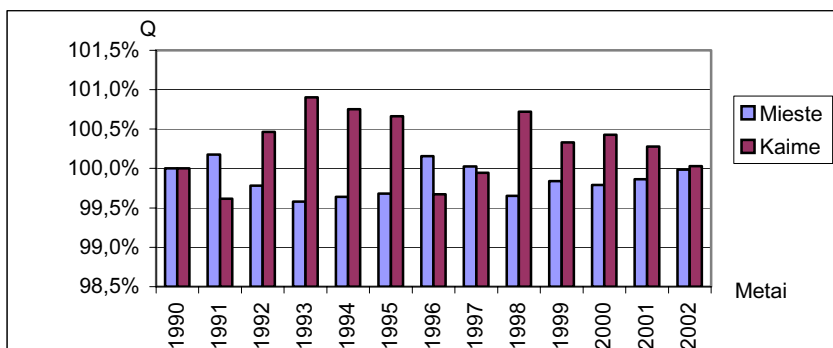
3.1 pav. Vidutinis metinis miesto ir kaimo gyventojų skaičiaus santykinis kitimo greitis G , apskaičiuotas pagal (3.6) formulę

Kitimo greitį galima apskaičiuoti ir antruoju būdu: nagrinėjama reikšmė palyginama su prieš tai buvusia reikšme. Šiuo atveju pagal (3.7)

formulę gaunamas santykinis kitimo greitis Q parodo, kiek pakito miesto ir kaimo gyventojų skaičius palyginti su prieš tai buvusiais metais. Šiuo būdu apskaičiuotas santykinis kitimo greitis yra grandininis:

$$Q_i = \frac{y_i}{y_{i-1}} \cdot 100, \quad i = 1, 2, \dots, n. \quad (3.7)$$

Pastaba: jei šių formulių dešiniųjų pusių nedauginsime iš 100, didėjimo greitis bus reiškiamas ne procentais, o kitimo koeficientais.



3.2 pav. Vidutinis metinis miesto ir kaimo gyventojų skaičiaus santykinis kitimo greitis Q , apskaičiuotas pagal (3.7) formulę

3.5. Demografinių duomenų pradinės analizės pavyzdys

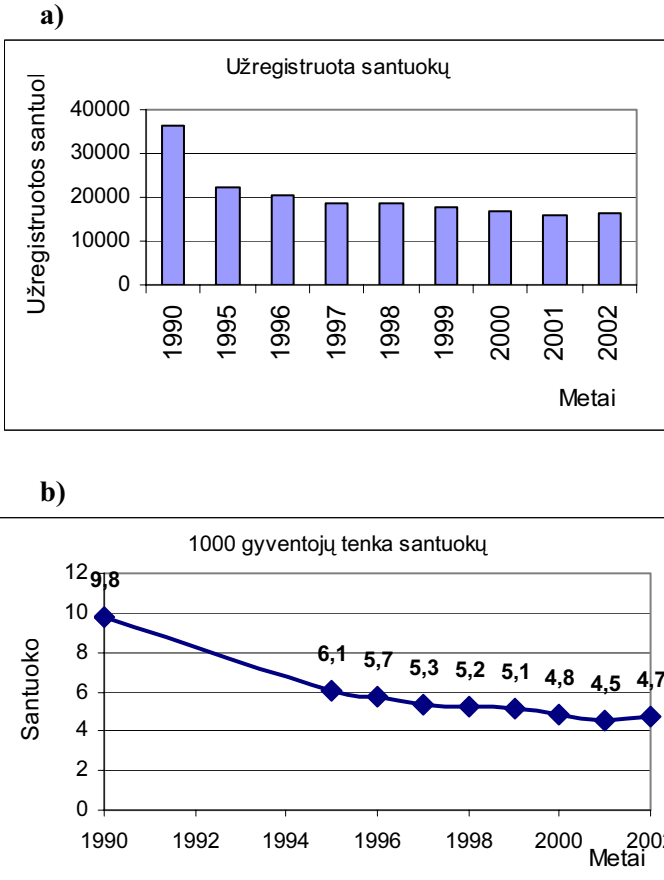
1990–2002 m. Lietuvoje užregistruotų santuokų skaičius pateiktas 3.3 lentelėje.

3.3 lentelė. Lietuvoje užregistruotos santuokos 1990–2002 m.

	1990	1995	1996	1997	1998	1999	2000	2001	2002
Užregistruota santuokų	36310	22150	20433	18796	18486	17868	16906	15764	16151
1000 gyventojų tenka santuokų	9,8	6,1	5,7	5,3	5,2	5,1	4,8	4,5	4,7

1. Duomenų diagramos

Duomenys dažniausiai pateikiami stulpelinėmis ir sklaidos (*scatter*) diagramomis (3.3 pav.).



3.3 pav. Rodiklio kitimo laike grafiniai vaizdai: a – stulpelinė diagrama, b – sklaidos diagrama

2. Pagrindinės tendencijos

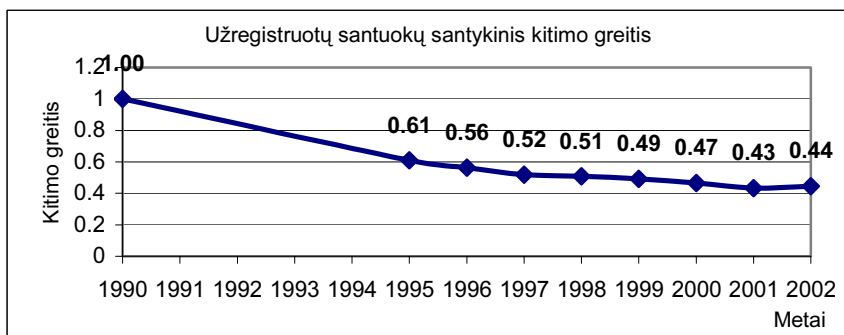
Iš pateiktų diagramų matyti, kad paskutinį dešimtmetį Lietuvoje užregistruotų santuokų skaičius nuolat mažėjo. Absoliutus santuokų skai-

čius sumažėjo nuo 36,3 tūkst. 1990 m. iki 16,2 tūkst. 2002 m. Užregistruotų 2002 m. santuokų skaičius sudarė tik 44 procentus santuokų, užregistruotų 1990 m. ($16151/36310=0,44$). Bendros proceso kitimo tendencijos išryškėja apskaičiuavus užregistruotų santuokų santykinį kitimo greitį (3.4 lentelė).

3.4 lentelė. Užregistruotų santuokų santykinis kitimo greitis palyginti su 1990 m.

Metai	1990	1995	1996	1997	1998	1999	2000	2001	2002
Santuokų kitimo greitis	1	0,610	0,563	0,518	0,509	0,492	0,466	0,434	0,445

Santykinis kitimo greitis atskleidžia, kaip keičiasi šis procesas palyginti su 1990 m. Matome, kad nuo 1997 m. užregistruotų santuokų skaičius mažėjo lėčiau (3.4 pav.), o 2002 m. pirmą kartą po ilgo laikotarpio netgi truputį padidėjo.



3.4 pav. Užregistruotų santuokų skaičiaus santykinis kitimas palyginti su 1990 m.

3.5 lentelėje apskaičiuotas 1990–1995 m. užregistruotų santuokų santykinis grandininis kitimo greitis (palyginti su prieš tai buvusiais metais).

3.5 lentelės duomenys teigia, kad užregistruotų santuokų labiausiai mažėjo 1990 – 1995 m. Šiuo laikotarpiu, palyginti su 1990 m., kasmet vidutiniškai sumažėdavo 0,74 santuokų 1000 gyventojų ($((6.1-9.8)/5)$). 1996–1997 m. vidutiniškai kasmet sumažėdavo 0,4 santuokų 1000 gy-

ventojų (palyginti su prieš tai buvusiais metais), 1998–1999 m. sumažėdavo 0,1 santuokų ir 2000–2001 m. 0,3 santuokų. 1996–2002 metų laikotarpiu kasmet vidutiniškai sumažėdavo 0,2 santuokų 1000 gyventojų.

3.5 lentelė. Užregistruotų santuokų 1000 gyventojų grandininis kitimo greitis 1990–1995 m.

	1995	1996	1997	1998	1999	2000	2001	2002
Užregistruotų santuokų skaičiaus santykinis grandininis kitimo greitis	-3,7*	-0,4	-0,4	-0,1	-0,1	-0,3	-0,3	0,2

* 1995 m. pokytis apskaičiuotas pagal 1990 m.

Užduotys

3.1. Jungtinių Tautų duomenys apie vidutinį vaikų skaičių, tenkanti vienai moteriai, pateikti 3.6 lentelėje. Apskaičiuokite vidutinį ir santykinį vaikų skaičius kitimą valstybėse ir suformuluokite išvadas apie 1975–1980 ir 1995–2000 m. vaisingumo rodiklio kitimą pasaulyje.

3.6 lentelė. Įvairių pasaulio valstybių suminiai vaisingumo rodikliai 1975–1980 ir 1995–2000 metais

<i>Suminis vaisingumo rodiklis (vidutinis vienos moters vaikų skaičius)</i>								
AZIJA	1975-1980	1995-2000	EUROPA	1975-1980	1995-2000	EUROPA	1975-1980	1995-2000
Armėnija	2,5	1,7	Albanija	4,2	2,5	Slovakija	2,50	1,40
Azerbaidžanas	3,6	2,0	Austrija	1,6	1,4	Slovėnija	2,20	1,30
Bangladešas	6,7	3,1	Baltarusija	2,10	1,40	Ispanija	2,60	1,20
Butanas	5,9	5,5	Belgija	1,70	1,60	Švedija	1,60	1,60
Kambodža	4,1	4,6	Bosnija ir Hercegovina	2,20	1,40	Šveicarija	1,50	1,50
Kinija	3,3	1,8	Bulgarija	2,20	1,20	Ukraina	2,00	1,40
Gruzija	2,4	1,9	Kroatija	2,00	1,60	Jungtinė Karalystė	1,70	1,70
Indija	4,8	3,1	Čekija	2,30	1,20	Jugoslavija	2,40	1,80

<i>Suminis vaisingumo rodiklis (vidutinis vienos moters vaikų skaičius)</i>								
AZIJA	1975-1980	1995-2000	EUROPA	1975-1980	1995-2000	PIETRYČIŲ AZIJA IR Š. AFRIKA	1975-1980	1995-2000
Indonezija	4,7	2,6	Danija	1,70	1,70			
Japonija	1,8	1,4	Estija	2,10	1,30	Afganistanas	7,20	6,90
Kazachstanas	3,1	2,3	Suomija	1,60	1,70	Alžyras	7,20	3,80
Šiaurės Korėja	3,3	2,1	Prancūzija	1,90	1,70	Egiptas	5,30	3,40
Pietų Korėja	2,9	1,7	Vokietija	1,50	1,30	Iranas	6,50	2,80
Kirgizija	4,1	3,2	Graikija	2,30	1,30	Irakas	6,60	5,30
Laosas	6,7	5,8	Vengrija	2,10	1,40	Izraelis	3,40	2,70
Malaizija	4,2	3,2	Islandija	2,70	2,10	Jordanas	7,40	4,90
Mongolija	6,6	2,6	Airija	2,40	1,90	Kuveitas	5,90	2,90
Mianmaras	5,3	2,4	Italija	1,60	1,20	Libanas	4,30	2,70
Nepalas	6,2	4,5	Latvija	1,80	1,30	Libija	7,40	3,80
Pakistanas	7,0	5,0	Lietuva	2,30	1,40	Marokas	5,90	3,10
Filipinai	5,0	3,6	Makedonija	2,70	2,10	Omanas	7,20	5,90
Singapūras	1,9	1,7	Moldavija	2,40	1,80	Saudo Arabija	7,30	5,80
Šri Lanka	3,8	2,1	Olandija	1,60	1,50	Sirija	7,40	4,00
Tadžikistanas	5,9	4,2	Norvegija	1,80	1,90	Tunisas	5,70	2,60
Tailandas	4,3	1,7	Lenkija	2,30	1,50	Turkija	4,50	2,50
Turkmėnija	5,3	3,6	Portugalija	2,40	1,40	Jungtiniai Arabų Emyratai	5,70	3,40
Uzbekistanas	5,1	3,4	Rumunija	2,60	1,20	Jemenas	7,60	7,60
Vietnamas	5,6	2,6	Rusija	1,90	1,30			

Šaltinis: United Nations Population Division

3.2. 3.7 lentelėje pateikti Lietuvos statistikos departamento duomenys apie ištuokų skaičių. Nubraižykite reikalingas diagramas, išanalizuokite ištuokų skaičiaus kitimo tendencijas ir apskaičiuokite padėties charakteristikas. Gautus rezultatus palyginkite su užregistruotų santuokų kitimu.

3.7 lentelė. Lietuvoje užregistruotos ištuokos 1990–2002 m.

	1990	1995	1996	1997	1998	1999	2000	2001	2002
Užregistruota ištuokų	12747	10221	11311	11371	11752	11390	10882	11024	10579
1000 gyventojų tenka ištuokų	3,4	2,8	3,1	3,2	3,3	3,2	3,1	3,2	3,1

3.3. Naudodamiesi Lietuvos statistikos departamento pateiktais duomenimis (3.8 lentelė), atlikite miesto ir kaimo gyventojų struktūros kitimo analizę. Taikydami santykinius dydžius, pavaizduokite šį kitimą diagramomis ir suformuluokite išvadas apie 1990–2002 m. vykusį vyrų ir moterų skaičiaus kitimą mieste ir kaime.

3.8 lentelė. Vidutinis metinis gyventojų skaičius tūkstančiais

Metai	Mieste		Kaime	
	Vyrai	Moterys	Vyrai	Moterys
1990	1186,5	1333,8	562,7	614,8
1991	1190,1	1339	561,7	613,3
1992	1186,1	1334,8	563,3	615,9
1993	1174	1324,4	565,7	618,5
1994	1158,9	1313,4	566,4	618,5
1995	1142,8	1302,7	566,6	617
1996	1132,1	1298,7	561,6	609,2
1997	1120,4	1293,2	558,4	603,2
1998	1105,1	1282,7	559,5	602
1999	1092,6	1274,5	558,3	598,8
2000	1079,9	1265,7	557,7	596,2
2001	1070,9	1259,3	556,8	594,3
2002	1065,9	1255,8	555	592,4

3.4. 3.9 lentelėje pateikti duomenys apie emigraciją iš Lietuvos 1928–1939 metais. Kaip kito tautinė emigrantų sudėtis? Kurią dalį išvykusiųjų sudarė lietuviai, kurią – kitų tautybių gyventojai? Išvadas iliustruokite diagramomis.

3.9 lentelė. Emigracija iš Lietuvos 1928–1939 metais ir emigrantų tautybė

Metai	Iš viso	Emigrantų tautybė						
		Lietuviai	Žydai	Rusai	Vokiečiai	Lenkai	Kiti	Užsieniečiai
Iš viso	41268	24327	13898	1320	536	190	77	920
1928	7251	5029	1664	395	106	44	13	–
1929	15999	12030	2825	457	267	105	39	276
1930	6428	4185	1736	64	104	19	10	310
1931	1756	476	1098	5	26	3	6	142
1932	1001	222	717	9	4	2	5	42
1933	1300	251	1020	3	2	–	1	23
1934	1521	345	1101	33	10	7	–	25
1935	1911	367	1418	88	4	–	–	34
1936	1707	407	1007	252	3	2	2	34
1937	979	487	447	3	6	2	–	34
1938	811	366	426	10	3	6	–	–
1939	604	162	439	1	1	–	1	–

3.5. Naudodamiesi Lietuvos statistikos departamento duomenimis apie gyventojų ir gimusių vaikų skaičių (3.10 lentelė) apskaičiuokite, kiek tenka gimusių vaikų 1000 gyventojų. Rezultatus pavaizduokite grafiškai ir parašykite išvadas apie gimusių vaikų skaičiaus kitimą 1990 – 2002 m.

3.10 lentelė. Natūralus gyventojų judėjimas

Metai	1990	1995	1996	1997	1998	1999	2000	2001	2002
Gimusių vaikų skaičius	56868	41195	39066	37812	37019	36415	34149	31546	30014
Gyventojų skaičius tūkst.	3697,8	3629,1	3601,6	3575,2	3549,3	3524,2	3499,5	3481,3	3469,1

3.6. Demografinės prognozės modeliai

Visų lygių valdymo procesuose vienas iš svarbiausių prognozės uždavinių yra įvertinti artimesnius ir tolimesnius gyventojų skaičiaus ir jų sudėties pokyčius.

Prieš pradėdant demografinės prognozės metodų analizę reikėtų aptarti keletą prognozėje vartojamų sąvokų (arba šios srities terminiją). Būsimiems demografiniams pokyčiams įvertinti vartojamos trys sąvokos: įvertis (arba įvertinimas), planavimas ir prognozė.

Įvertis – tai netiesioginis matas sąlygų ar rodiklių, kurių negalima išmatuoti tiesiogiai. Pavyzdžiui: įvairių autobusų maršrutų krūvį galima įvertinti per keletą pasirinktų dienų skaičiuojant, kiek žmonių važiuoja autobusais.

Įverčiams būdinga tai, kad vertinamos tik ankstesnės arba esamos sąlygos, ir nevertinamos būsimos. Šio tipo prognozė remiasi jau žinomomis sąlygomis ir stengiasi kuo tiksliau įvertinti tiriamas sąlygas ar rodiklius.

Planavimas yra sąlyginis „jeigu, tai“ (*if-then*) teiginys apie ateitį. Pavyzdžiui, jeigu nuo 1990 iki 2000 m. miestas išaugo 10 procentų, t. y. nuo 10 000 gyventojų iki 11 000, tai, išlikus tokiam pat augimo greičiui, 2010 m. šiame mieste turėtų gyventi 12 100 gyventojų. Planavimas yra populiariausias metodas, kadangi, jei skaičiavimai yra teisingi, neteisingos išvados padaryti beveik neįmanoma. Planavimas paprasčiausiai teigia: „Jei atsitiks taip, tai ateitis bus tokia. Jeigu atsitiks kitaip, tai ir šio rezultato nebus.“

Prognozė – tai sudėtingesnis nei planavimas labiausiai tikėtinos ateities numatymo metodas. Pavyzdžiui: turėdamas visus galimus planus, analitikas gali nutarti, kad dėl didelio ekonomikos augimo ir didėjančio miesto populiarumo nuo 2000 iki 2010 m. gyventojų jame turėtų padidėti ne 10, o 20 procentų. Dėl didelio sudėtingumo, galimo netikslumo ir galimų prieštaringų rezultatų prognozę dažniausiai atlieka šios srities specialistai. Deja, prognozės retai būna labai tikslios, todėl dažniausiai ne prognozuojama, o planuojama.

Būsimoms sąlygoms planuoti taikomi dviejų pagrindinių grupių metodai: **aprosimacijos** metodai ir **kohortų komponentiniai** metodai.

Prieš nagrinėdami aprosimacijos metodus supažindinsime su laiko eilučių ir statistinio ryšio sąvokomis.

3.7. Laiko eilutės ir aproksimacijos metodai

Laiko eilutės sąvoka. Demografijoje aproksimacijos metodai dažniausiai taikomi *laiko eilutėms*. Laiko eilutė – tai skaičių seka, kurios elementai – stebimo proceso laike reikšmės. Šios reikšmės matuojamos nuosekliais laiko momentais, vienodais laiko intervalais. Paprastai laiko eilutės elementai numeruojami pagal laiko momentą, kuriuo jie išmatuoti (x_1, x_2, x_3 ir t.t.). Todėl eilė, kuria surašytos išmatuotos reikšmės, yra labai svarbi.

Platesnė laiko eilutės sąvoka. Dažnai laiko eilutė yra suprantama plačiau. Pavyzdžiui, vienu metu gali būti registruojama keletas nagrinėjamo reiškinio charakteristikų. Šiuo atveju kalbama apie daugiamates arba daugialypes laiko eilutes. Jeigu matavimai vyksta nenutrūkstamai, kalbama apie tolydines laiko eilutes, arba *atsitiktinius procesus*. Pagaliau tiriamas reiškinys gali kisti ne laike, o koku nors kitu būdu, pavyzdžiui, erdvėje. Tada kalbama apie atsitiktinius laukus.

Laiko eilučių kitimo forma gali būti labai įvairi. 3.5 pav. pateikta įvairių laiko eilučių pavyzdžių.

Laiko eilučių analizės tikslas. Praktiškai tiriant laiko eilutes apsiribojama tam tikros laiko atkarpos stebėjimais. Remiantis šiais stebėjimais, reikia padaryti išvadas apie šios stebimos eilutės savybes ir tikimybiškai įvertinti šio proceso priežastis. Nagrinėjant laiko eilutes, dažniausiai keliami šie tikslai:

- eilutei būdingų savybių ir ypatybių nustatymas;
- statistinio modelio, tinkančio laiko eilutei, parinkimas;
- būsimų reikšmių prognozavimas pagal ankstesnes proceso reikšmes.

Praktiškai šie tikslai pasiekiami ne visada ir nevysiškai. Dažnai trukdo nepakankama stebėjimų apimtis (trumpos sekos), o dar dažniau – laike besikeičianti nagrinėjamos laiko eilutės struktūra. Dėl tokio kitimo ankstesnės stebėjimų reikšmės tampa bevertės ir jos nepadeda numatyti ateities.

Laiko eilučių tyrimo etapai:

- grafinis laiko eilutės pateikimas ir kitimo aprašymas;
- dėsningai kintančios laiko eilutės dedamosios (trendo) išskyrimas bei matematinio modelio sudarymas;
- laiko eilute aprašomo proceso būsimų reikšmių prognozavimas;
- ryšių tarp atskirų laiko eilučių nustatymas.

Pagal pagrindinę statistikos idėją, analizuojant laiko eilutę, jos kintamą stengiamasi padalinti į dvi dalis: dėsningai kintančią (trendą) ir atsitiktinę. Dėsningai kintanti laiko eilutės dalis (dedamoji) yra pavaldi tam tikram dėsniui, ir todėl prognozuojama. Tokią dedamąją x_t galima apskaičiuoti kiekvienu laiko momentu t kaip šio laiko momento t funkciją. Ši funkcija, be laiko t , dar gali priklausyti nuo tam tikrų parametrų. Jeigu parametrai nežinomi, juos tenka įvertinti vadovaujantis turimų stebėjimų reikšmėmis.

Aproksimacijos metodai. Dėsningai kintančios laiko eilutės dalies išskyrimo metodai demografijoje dažnai vadinami aproksimacijos metodais. Aproksimacijos metodai – tai „kreivės pritaikymo“ metodai, kai paskutiniųjų metų duomenims parenkama geriausiai tinkanti kreivė, ir ši kreivė pratęsiama būsimums reikšmėms. Tokia nesudėtinga procedūra dažnai naudojama planuojant atskirų vietovių demografinius pokyčius tais atvejais, kai nėra visų reikiamų duomenų apie kitas sritis. Čia sudėtingiausia yra parinkti nagrinėjamų metų skaičių. Kai kitimas yra lėtas, tyrimui naudingus duomenis parinkti nesudėtinga, nes jiems tinkanti kreivė pasižymi lėtu kitimu. Tačiau nagrinėjant greitai augančius ar mažėjančius rajonus ir norint parinkti atitinkamą demografinę kreivę, kelių dešimtmečių gyventojų skaičiaus nagrinėjimas bus mažai naudingas, o taikant šį metodą labiau tiktų naujesni duomenys.

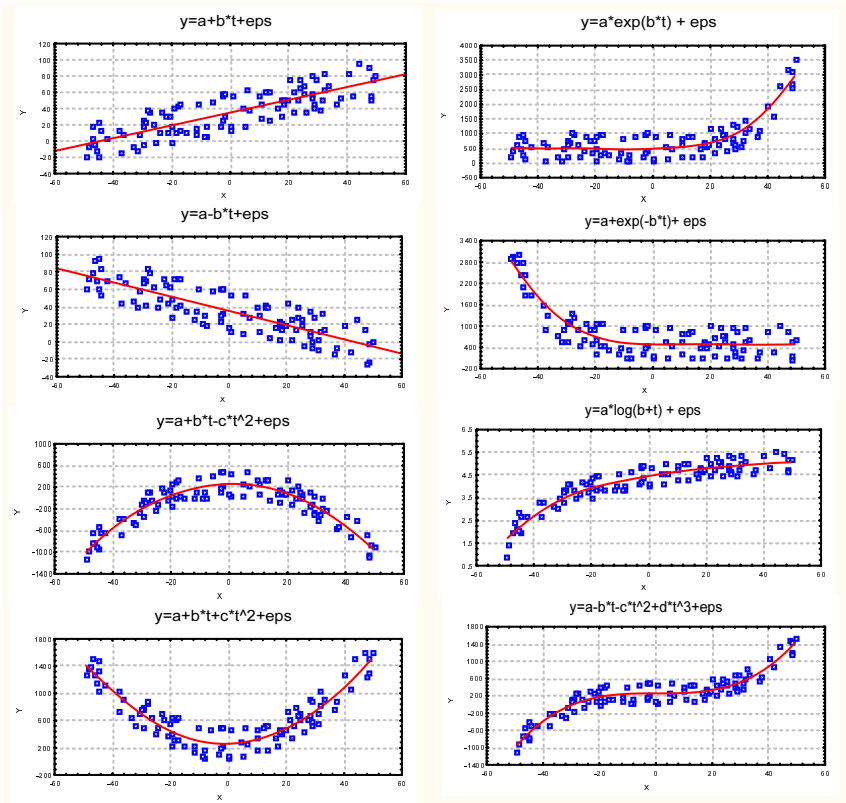
Šio metodo privalumas yra jo paprastumas, o trūkumas – naudojamos tik to paties proceso būtosios reikšmės, neįvertinant jokių kitų faktorių – gyvenamųjų sąlygų kitimo, ekonominių pokyčių ir pan. Kitas trūkumas yra tas, kad ateitis planuojama pagal būtašias sąlygas manant, kad praeities sąlygos padeda planuoti ateitį. Tačiau nėra jokių garantijų, kad praeitis turės didelės įtakos ateičiai, todėl nėra tikrumo, kad praeities trendas tęsis į ateitį. Dėl to šį metodą reikia taikyti atsargiai ir tik tais atvejais, kai nėra galimybių ar labai sudėtinga taikyti kitus, tobulesnius, metodus.

Taikant aproksimacijos metodą tenka spręsti tris problemas:

1. Parinkti tinkamą aproksimacijos funkciją;
2. Įvertinti aproksimacijos paklaidą;
3. Parinkti tinkamo ilgio seką.

Per duotus taškus galima nubrėžti daug funkcijų, todėl pirmasis uždavinys – parinkti geriausiai tinkančią. Demografijoje naudojamos įvairių rūšių funkcijos – nuo paprasčiausių iki labai sudėtingų. Prieš parenkant funkciją reikia nustatyti, kokio tipo priklausomybė sieja kintamuo-

sius, t. y. pasirinkti funkcijos tipą. Šiuo tikslu braižoma nagrinėjamų duomenų diagrama. Šia diagrama nustatoma, kokio pavidalo matematinė funkcija geriausiai turėtų tikti. Dažniausiai pasirenkamos funkcijų formos pavaizduotos 3.5 pav.



3.5 pav. Kintamojo y formos ir jas aprašančios matematinės funkcijos $y=f(t)$

Nagrinėjami duomenys retai kada išsidėsto tiksliai tiesiąja linija ar kreive, dažniausiai jie telkiasi apie šią kreivę, todėl įvedama atsitiktinė paklaida ε (diagramose pažymėta *eps*). Paprasčiausia tiesinė funkcija, įvedus atsitiktinę paklaidą, užrašoma lygtimi:

$$y=a+bt+\varepsilon. \quad (3.8)$$

Per turimus taškus galima nubrėžti daug tiesiųjų linijų, todėl iškyla problema pasirinkti pačią tinkamiausią. Galima tiesiųjų linijų realizacija priklausyti nuo atsitiktinės paklaidos ε , todėl yra nustatyti reikalavimai, kuriuos turi tenkinti šios paklaidos. Reikalavimai apibrėžiami taip:

1. Paklaidą ε sudarantys atsitiktiniai dydžiai turi normalųjį skirstinį;
2. ε vidurkis lygūs nuliui;
3. Visų ε dispersijos yra vienodo dydžio;
4. Visi ε yra nepriklausomi.

Parenkant funkciją reikia atsižvelgti ne vien į diagramos pavidalą, bet ir į tai, kokių tikslų funkcija bus naudojama. Pavyzdžiui, demografi-koje dažnai naudojama eksponentinė funkcija gerai tinka tada, kai reikia įvertinti žmonių ar kitų biologinių populiacijų augimą. Ši funkcija gerai tinka tada, kai reikia apskaičiuoti tam tikro istorinio laikotarpio vidutinį populiacijos augimo koeficientą. Tačiau ji blogai tinka paskutiniųjų trijų amžių žmonių populiacijos augimo prognozei. Apsiribosime šiomis funkcijomis:

- tiesine,
- logaritmine,
- polinome,
- laipsnine,
- eksponentine,
- logistine,
- Hompereso.

Tiesinė aproksimacija vaizduojama tiese, kurią galima užrašyti lygtimi: $y(t) = at + b + \varepsilon$; čia: a ir b yra konstantos (a vadinama pasvirimo koeficientu, o b – atkirtimo koeficientu).

Logaritminei aproksimacijai taikoma lygtis $y(t) = c \ln(t) + b + \varepsilon$; čia: c ir b yra konstantos, \ln – natūrinis logaritmas.

Polinomei aproksimacijai naudojamas n -tosios eilės polinomas:

$$y(t) = b + c_1 t + c_2 t^2 + c_3 t^3 + \dots + c_n t^n + \varepsilon; \text{ čia: } b \text{ ir } c_1, c_2, c_3, \dots$$

yra konstantos. Praktiniuose uždaviniuose polinomo eilė n labai retai kada viršija 5.

Laipsninei aproksimacijai naudojama laipsninė funkcija:

$$y(t) = ct^b + \varepsilon; \text{ čia: } c \text{ ir } b \text{ yra konstantos.}$$

Eksponentinei aproksimacijai taikoma lygtis: $y(t) = ce^{bt} + \varepsilon$; čia: c ir b yra konstantos, e – iracionalusis skaičius ($e=2,71828$). Šis modelis

dažniausiai naudojamas tada, kai laiko eilutė turi tendenciją išsaugoti pastovų kitimo greitį.

$$\text{Logistinio modelio lygtis: } y = \frac{a}{1 + b \cdot e^{-ct}} + \varepsilon.$$

$$\text{Hompertso modelis: } \log(y) = a - b \cdot r^t + \varepsilon, \text{ čia } 0 < r < 1.$$

Paskutiniai du modeliai gali tikti S formos funkcijai. Šio tipo funkcija sutinkama procesuose, kai pradinėje stadijoje proceso augimo greitis nuolat didėja, o vėliau mažėja. Tokių modelių panaudojimą lemia kai kurių socialinių bei ekonominių procesų savybė ilgą laiką vystytis vienu augimo greičiu. Tokių procesų aproksimacijai tinka ir polinominiai modeliai.

Pirmuoju žingsniu, parenkant galimos aproksimacijos funkcijos pavidalą, piešiamas grafinis duomenų vaizdas. Prognozavimui aproksimacinės funkcijos gerai tinka tuo atveju, kai stebima ilga proceso atkarpa. Kai seka trumpa, prognozė dažnai būna netiksli.

Smulkiau apie rodiklių ryšių stiprumo nustatymą ir aproksimacijos funkcijos parametrų vertinimą skaitykite skyriuje „Teisinė statistika“.

3.8. Aproksimacijos metodo taikymo pavyzdys

Panagrinėkime Lietuvos statistikos departamento duomenis apie gyvus gimusius kūdikius 1995–2002 metais (3.11 lentelė). Naudodamiesi aproksimacijos lygtimi, apskaičiuokime tikėtiną vaikų skaičių 2003 m.

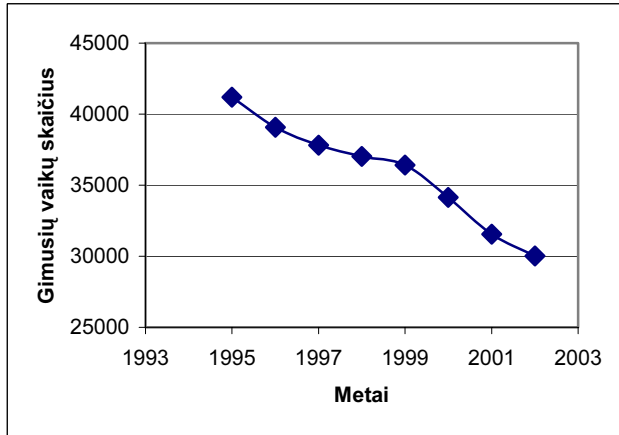
3.11 lentelė. 1995–2002 m. Lietuvoje gimusių gyvų kūdikių skaičius

Metai (t)	1995 (1)	1996 (2)	1997 (3)	1998 (4)	1999 (5)	2000 (6)	2001 (7)	2002 (8)
Gimusių gyvų kūdikių skaičius (y)	41195	39066	37812	37019	36415	34149	31546	30014

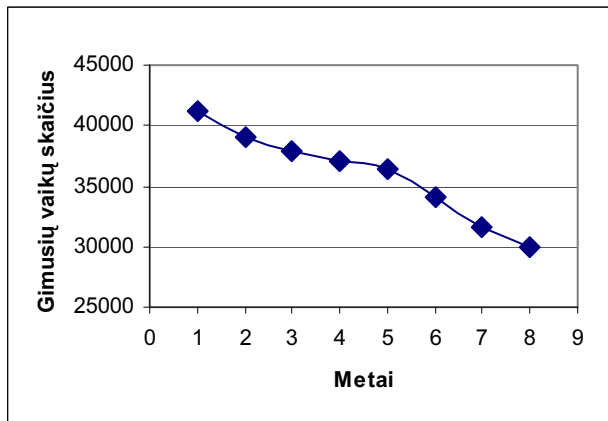
Pirmiausia šiuos duomenis pavaizduokime grafiškai – nubraižykime jų sklaidos diagramą (3.6 pav.). 3.6 a diagramoje horizontaliojoje X ašyje atidėtos tikrosios metų reikšmės. 3.6 b diagramos X ašyje atidėtos

santykinių metų reikšmės (1,2,...). Santykinės metų reikšmes skaičiavimams patogiau naudoti nei tikrąsias, kadangi gaunamos patogesnės (mažesnės) koeficientų reikšmės.

a)



b)



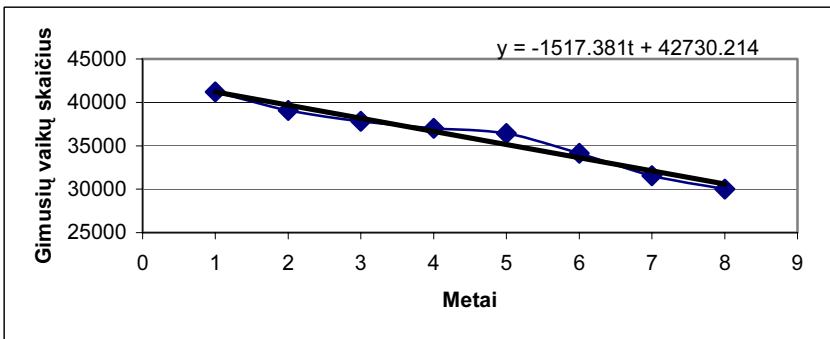
3.6 pav. 1995–2002 m. Lietuvoje gimusių gyvų kūdikių skaičiaus sklaidos diagramos (a – realus laikas, b – santykinis)

Kaip matyti iš 3.6 pav., gimusių gyvų kūdikių skaičius nuo 1995 iki 2002 m. nuolat mažėjo, todėl jo kitimui prognozuoti galima naudoti tiesinę funkciją.

Parinkus aproksimacijos funkcijos (tiesės ar kreivės) tipą, koeficientai skaičiuojami kompiuteriu, pasitelkus tinkamą programinę įrangą. Šiuo tikslu gali būti naudojamos įvairios programos, pradedant nuo skaičiuoklių ir baigiant sudėtingomis statistinėmis sistemomis.

Apskaičiavus koeficientus gaunama lygtis (3.9), pagal kurią nubrėžiama aproksimacinė tiesė (3.7 pav.):

$$y = -1517,381 * t + 42730,214 . \quad (3.9)$$



3.7 pav. 1995–2002 m. Lietuvoje gimusių gyvų kūdikių skaičiaus sklaidos diagrama su aproksimacine (trends) linija

Remiantis šia lygtimi, galima prognozuoti kūdikių skaičių artimiausiais metais, jei „gimdymo rizika“ pasižyminčių moterų socialinės bei ekonominės sąlygos išliks nepakitusios:

$$y = -1517,381 * 9 + 42730,214 \approx 29074 .$$

Taigi galima tikėtis, kad nesikeičiant sąlygoms, 2003 metais gims 29074 naujagimių.

Užduotys

3.6. 3.12 lentelėje pateikti duomenys apie santuokas, emigraciją, gimstamumą ir mirtingumą 1000 gyventojų 1920–1939 metais. Santuokų, gimimų ir mirčių skaičiaus prognozei parinkite tinkamiausią aproksimacijos lygties pavidalą ir nustatykite 1940 m. tikėtiną santuokų, gimimų ir mirčių skaičių.

3.12 lentelė. Santuokos, emigracija, gimstamumas ir mirtingumas 1000 gyventojų 1920–1939 metais

Metai	Santuokos	Gimė	Mirė	Prieauglis	Emigravo	Kūdikių iki 1 m. mirtingumas %
1920	8,2	22,7	21,2	1,5
1921	7,4	24,6	15,2	9,4
1922	7,7	27,3	17,7	9,6	...	16,9
1923	7,1	28,4	15,1	13,3	1,3	14,8
1924	8,1	29,4	16,3	13,1	1,4	17
1925	7,9	28,9	16,9	12	1,3	17,9
1926	8,5	28,5	15,4	13,1	4,6	14,6
1927	8,6	29,4	17,3	12,1	8,0	15,1
1928	8,3	28,8	15,6	13,2	3,7	14,7
1929	7,5	27,2	17,1	10,1	6,9	17,6
1930	7,8	27,4	15,9	11,5	2,7	15,4
1931	8,0	26,8	15,8	11,0	0,7	14,5
1932	7,9	27,3	15,3	12,0	0,4	16,7
1933	8,1	25,7	13,5	12,1	0,5	12,1
1934	7,4	24,8	14,6	10,2	0,6	16,6
1935	7,4	23,4	14,0	9,4	0,8	12,3
1936	7,5	24,2	13,4	10,8	0,7	12,8
1937	7,4	22,3	13,2	9,1	0,4	12,0
1938	7,9	22,7	12,6	10,1	0,3	11,3
1939	7,5	22,4	13,6	8,8	0,2	12,2

3.7. 3.13 lentelėje pateikti Lietuvos statistikos departamento duomenys apie tikėtiną gyvenimo trukmę gimstant. Atlikite miesto ir kaimo gyventojų tikėtinos gyvenimo trukmės kitimo analizę. Parinkite tinkamiausią aproksimacijos lygtį ir nustatykite 2001 m. vyrų ir moterų tikėtiną gyvenimo trukmę gimstant.

3.13 lentelė. Lietuvos gyventojų tikėtina gyvenimo trukmė gimstant

Metai	Iš viso			Mieste			Kaime		
	Iš viso	Vyrai	Moterys	Iš viso	Vyrai	Moterys	Iš viso	Vyrai	Moterys
1990	71,49	66,55	76,22	72,43	67,6	76,78	69,48	64,16	75,12
1991	70,71	65,28	76,07	71,78	66,55	76,67	68,49	62,65	74,98
1992	70,49	64,92	76,02	71,69	66,34	76,71	68,15	62,14	74,84
1993	69,06	63,26	75,01	70,32	64,7	75,81	66,59	60,48	73,62
1994	68,71	62,73	74,89	69,76	63,98	75,48	66,41	60,19	73,66
1995	69,29	63,53	75,15	70,37	64,92	75,64	67,03	60,79	74,28
1996	70,5	64,97	76	71,83	66,63	76,69	67,85	61,8	74,73
1997	71,41	65,9	76,82	72,82	67,61	77,65	68,63	62,64	75,32
1998	71,78	66,5	76,87	73,26	68,34	77,75	68,94	63,16	75,36
1999	72,33	67,07	77,41	74,09	69,15	78,58	69,15	63,45	75,47
2000	72,87	67,62	77,93	74,37	69,38	78,96	70,08	64,49	76,22

3.9. Kohortų (grupių) komponentiniai metodai

Ekstrapoliacijos metodų galimybės yra labai ribotos. Iš gauto skaičiaus, kuris yra apibendrintas gyventojų populiacijos rodiklis, nieko negalima pasakyti apie tai, kaip keisis skirtingo amžius gyventojų grupių dydis bei vyrų ir moterų skaičius. Ar daugės senų žmonių, ar daugės vaikų? Ar daugės jaunų žmonių, galinčių turėti daugiau vaikų? Net jei bendras gyventojų skaičius nekinta, populiacijos sudėtis, ypač pagal amžiaus grupes, yra labai svarbi planuojant viešąją infrastruktūrą ir paslaugas. Vaikų daugėjimas reiškia, kad reikia planuoti daugiau mokyklų ir su tuo susijusių paslaugų. Suaugusių jaunų gyventojų grupių didėjimas taip pat turi didžiulę įtaką, kadangi šio amžiaus gyventojai daug važinėja, auginą vaikus, ir jiems reikia daugiau paslaugų nei kitų amžių grupėms. Senyvo amžiaus gyventojams reikia dar kitokių paslaugų. Tai gi norint gauti detalesnės informacijos apie esamą ir būsimą gyventojų populiaciją, į planavimo ar prognozavimo metodus būtina įtraukti amžių. Be amžiaus, į pažangesnių ir sudėtingesnių metodų lygtis įtraukiama ir lytis, kadangi vyrų ir moterų mirtingumas amžiaus grupėse yra skirtingas, be to, tik viena iš šių grupių, moterys, „rizikuoja“ turėti vaikų.

Nuodugnesnio ir tikslesnio planavimo poreikių patenkinti buvo sukurti kohortų komponentiniai planavimo metodai. Šiais metodais aprėpiama informacija apie amžiaus/lyties kohortas ir prognozuojama, kaip keisis jų dydis keičiantis visai populiacijai.

Populiacijos kohortos

Kaip jau buvo rašyta, gyventojų pasiskirstymas pagal amžių yra labai svarbus rodiklis planuojant infrastruktūrą ir paslaugas. Amžiaus įtraukimas į modelį atliekamas suskirstant gyventojų populiaciją į amžiaus kohortas. Dažniausiai naudojamas amžiaus intervalas yra penkeri metai. Retsykiais naudojamas ir dešimties metų intervalas. Pavyzdžiui, Lietuvos statistikos departamentas pateikia tokius duomenis apie 2003 m. gyventojų amžiaus kohortas (3.14 lentelė):

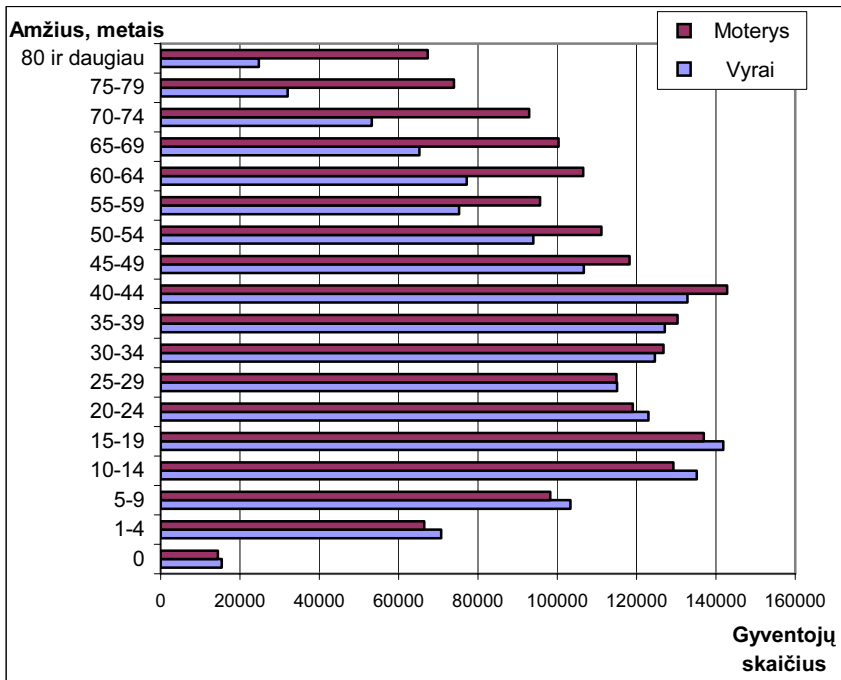
3.14 lentelė. Lietuvos gyventojų kohortos pagal lytį ir amžių

Lietuvos gyventojų skaičius pagal lytį ir amžių 2003 metų pradžioje			
Amžius	Iš viso	Vyrai	Moterys
0	29813	15379	14434
1–4	137159	70693	66466
5–9	201554	103307	98247
10–14	264439	135133	129306
15–19	278778	141877	136901
20–24	242033	122983	119050
25–29	230042	115097	114945
30–34	251389	124589	126800
35–39	257463	127106	130357
40–44	275645	132838	142807
45–49	224937	106716	118221
50–54	205044	93934	111110
55–59	170918	75281	95637
60–64	183654	77156	106498
65–69	165572	65259	100313
70–74	146131	53194	92937
75–79	105892	31979	73913
80+	92090	24783	67307
Iš viso	3462553	1617304	1845249

Šios penkerių metų gyventojų kohortos išsiskiria skirtingais mirtinumo rodikliais (mirtingumas didesnis pirmaisiais gyvenimo metais ir vyresniame mažiujė). Vyrų ir moterų mirtingumo rodikliai taip pat skiriasi. Vyrų mirtingumo rodiklis visose amžiaus kohortose išlieka didesnis nei moterų. Populiacija gali būti skirstoma ir pagal rasę (baltieji, juodieji, azijiečiai ir pan.), kadangi mažumą sudarančių rasių mirtingumo rodikliai dažniausiai būna didesni nei pagrindinės rasės gyventojų.

Amžiaus/lyties kohortos grafiškai vaizduojamos populiacijos piramide. Pavyzdžiui, Lietuvos gyventojų kohortų pagal lytį ir amžių 2003 metų pradžioje amžiaus/lyties piramidė pateikta 3.8 pav.

Amžiaus/lyties piramidėmis vaizdžiai perteikiama gyventojų struktūra pagal amžių ir deformacijos zonas. Iš 3.8 pav. matyti, kaip mažas vaikų iki 10 metų skaičius iškreipia piramidę. Pastebimas aiškus gimstamumo netolygumas: prieš mažą 1–4 m. amžiaus grupę stebima didelė 15–19 m. amžiaus grupė. Akivaizdu, kad švietimo įstaigoms teks siaurinti savo veiklą.



3.8 pav. 2003 m. pradžios Lietuvos gyventojų amžiaus/lyties piramidė

Akivaizdu, jog yra neatitikimas tarp vyresnio amžiaus vyrų ir moterų skaičiaus. Vyresnio amžiaus vyrų skaičius mažėja daug greičiau nei tokio pat amžiaus moterų skaičius.

Gyventojų amžiaus/lyties kitimo tendencijas patogu analizuoti lyginant paskutiniojo surašymo duomenis su ankstesniųjų surašymų duomenimis.

3.10. Tikėtina gyvenimo trukmė ir amžiaus lentelės

Tikėtinos gyvenimo trukmės ir amžiaus lentelių skaičiavimams taikomi statistiniai išlikimo analizės (*survival analysis*) metodai. Šie metodai pirmiausia buvo sukurti ir taikomi medicinos ir biologijos moksluose, tačiau vėliau paplito ir plačiai naudojami socialiniuose ir ekonomikos moksluose, taip pat ir inžinerijoje. Išlikimo analizės metoduose naudojama „cenzūruoto kintamojo“ sąvoka. Cenzūruoti stebėjimai atsiranda tada, kai mus dominantis rezultatas yra tam tikro, iš esmės baigtinio, įvykio trukmė. Tokie įvykiai būdingi įvairioms tyrimų sritims. Pavyzdžiui, socialiniuose moksluose dažnai nagrinėjama vedybų trukmė, studentų, iškritusių iš aukštųjų mokyklų procentiniai santykiai ir pan. Ekonomikoje tiriama naujų produktų, pavyzdžiui, automobilių, „išlikimo“ trukmė. Kokybės kontrolėje įprasta tirti apkrautas detales (suirimo laiko analizė).

Tiesiausias būdas, kuriuo galima įvertinti elementų išlikimą, yra jų amžiaus lentelė. Amžiaus lentelė yra vienas iš seniausių išlikimo analizės metodų, ir jau beveik 200 metų, kai ji yra pagrindinis draudimo ekspertų įrankis. Ir šiuo metu amžiaus lentelė yra vienas iš pagrindinių demografinės analizės metodų. Iš amžiaus lentelės yra apskaičiuojama tikėtina gyvenimo trukmė.

Amžiaus lentelės idėja labai paprasta. Įsivaizduokime didelę grupę (kohortą) žmonių, kurie gimė tą pačią dieną. Jei tirsime šią kohortą nuo pat gimimo, galėsime užregistruoti, kiek asmenų numirė per metus ir kiek asmenų yra gyvų per kiekvieną gimtadienį, tarkime, amžiaus x . Santykis mirusiųjų per metus su pradiniu kohortos skaičiumi yra tikimybė numirti sulaukus x amžiaus. Kai visiems x šios tikimybės tampa žinomos, amžiaus lentelė yra užbaigta.

Praktikoje tokia „kohortos amžiaus lentelė“ yra retai naudojama, kadangi asmenis gali tekti tirti iki 100 metų, o rezultatų lentelė atspindė-

tų buvusias istorines sąlygas, kurios jau nebegali būti taikomos. Vietoje tokios lentelės naudojama tam tikro laikotarpio, arba dabartinė, amžiaus lentelė. Sudarant šią lentelę yra sumuojamas visų amžiaus asmenų mirštamumas per trumpą laikotarpį, dažniausiai per vienerius ar trejus metus. Suskaičiuojama šio trumpo laikotarpio kiekvieno x amžiaus mirties tikimybė $q(x)$. Tam dažnai naudojama reguliariai atliekamo gyventojų surašymo informacija. Apskaičiuotos tikimybės $q(x)$ yra pritaikomos hipotetinei kohortai, susidedančiai iš 100 000 asmenų, ir taip gaunama amžiaus lentelė.

Pagal laiko intervalų ilgį amžiaus lenteles galima padalinti į dvi grupes:

- *visoje amžiaus lentelėje* pateikiami kiekvienių metų duomenys nuo gimimo iki paskutiniųjų metų. Tokios lentelės pildymas yra sudėtingas ir reikalauja daug laiko. Lentelės tikslumas priklauso nuo demografinių įvykių, kuriuos užregistravo statistinės registracijos sistemos, tikslumo.
- *sutrumpintoje amžiaus lentelėje*, skirtingai negu visoje, duomenys suskirstomi į intervalus, kurių ilgis nuo 5 iki 10 m. Tokias lenteles patogiau sudaryti ir jomis naudotis.

Paprastai tiriami kohorta, kurios dydis yra 100 000. Pirmieji gyvenimo metai dažniausiai išskiriami į atskirą intervalą. Likusių intervalų ilgis dažniausiai būna vienodas.

Sutrumpintos amžiaus lentelės sudarymo pavyzdys

Panagrinėkime amžiaus lentelės sudarymo principus, naudodamiesi JAV, Arkanzaso valstijos moterų amžiaus lentelės duomenimis (3.15 lentelė).

Amžiaus lentelė sudaroma apskaičiuojant atskirų stulpelių reikšmes.

x : *pirmajame stulpelyje* (x) visada pateikiamos amžiaus grupės. Sutrumpintoje amžiaus lentelėje pateikiami intervalai $[x, x+n)$, čia: n – intervalo ilgis, x – pradinė intervalo reikšmė. Pirmoji amžiaus grupė – vienerių metų, tada antroji – ketverių metų, o visos likusios iki pat paskutiniosios – penkerių metų. Paskutinioji amžiaus grupė lieka atvira, ir ji kartais žymima x^+ , kartais $>x$.

n : *antrajame stulpelyje* n – amžiaus intervalo ilgis.

l_x : trečiajame stulpelyje (l_x) – žmonių skaičius iš pradinės 100 000 dydžio kohortos sulaukusių amžiaus intervalo x pradžios. Pavyzdžiui, 30–35 m. amžiaus kohortos reikšmė 97 440 reiškia, kad iš 100 000 naujagimių 97 440 asmenys išgyveno bent iki 30 metų. Kadangi mirštama palaipsniui nuo pat gimimo, tai didėjant amžiui, l_x mažėja, ir visada $l_x \geq l_{x+1}$.

${}_n d_x$: ketvirtajame stulpelyje (${}_n d_x$) apskaičiuojamas kiekvieno n amžiaus intervalo mirčių skaičius, esant pradinei 100 000 dydžio naujagimių kohortai. Pavyzdžiui, 30–34 m. amžiaus intervale numirė 621 asmuo iš 97 440 asmenų, sulaukusių 30 metų: ${}_n d_x = l_x - l_{x+n}$.

3.15 lentelė. Moterų amžiaus lentelė (Arkanzasas, 2000).

Metai	n	l_x	${}_n d_x$	${}_n q_x$	${}_n L_x$	T_x	e_x
1	2	3	4	5	6	7	8
<1	1	100000	862	0,00862	99569,0	7769274	77,7
1–5	4	99138	195	0,00197	396162,0	7669705	77,4
5–10	5	98943	122	0,00123	494410,0	7273543	73,5
10–15	5	98821	137	0,00139	493762,5	6779133	68,6
15–20	5	98684	359	0,00364	492522,5	6285370	63,7
20–25	5	98325	404	0,00411	490615,0	5792848	58,9
25–30	5	97921	481	0,00491	488402,5	5302233	54,1
30–35	5	97440	621	0,00637	485647,5	4813830	49,4
35–40	5	96819	746	0,00771	482230,0	4328183	44,7
40–45	5	96073	1003	0,01044	477857,5	3845953	40,0
45–50	5	95070	1424	0,01498	471790,0	3368095	35,4
50–55	5	93646	1956	0,02089	463340,0	2896305	30,9
55–60	5	91690	3192	0,03481	450470,0	2432965	26,5
60–65	5	88498	5143	0,05811	429632,5	1982495	22,4
65–70	5	83355	7700	0,09238	397525,0	1552863	18,6
70–75	5	75655	10099	0,13349	353027,5	1155338	15,3
75–80	5	65556	12614	0,19242	296245,0	802310	12,2
80–85	5	52942	15571	0,29411	225782,5	506065	9,6
>85	10	37371	37371	1,0	280282,5	280282,5	7,5

Jei intervalo ilgis vieneri metai, tai $d_x = l_x - l_{x+1}$.

Paskutiniosios amžiaus grupės mirčių skaičius yra toks pat, koks ir šio intervalo pradžioje gyvų žmonių skaičius.

${}_nq_x$: penktajame stulpelyje (${}_nq_x$) apskaičiuojama tikimybė numirti tarp x ir $x+n$ amžiaus. Šis rodiklis atspindi grupės $[x, x+n)$ vidutinį mirtingumo lygį. Reikšmė 0,0064 amžiaus intervale 30–35 metai reiškia, kad iš 10 000 asmenų, sulaukusių 30 metų, 64 numirs nesulaukę 35-ojo gimtadienio. Ši funkcija apskaičiuojama pagal formulę:

$${}_nq_x = \frac{{}_nd_x}{l_x}.$$

${}_nL_x$: šeštajame stulpelyje (${}_nL_x$) yra informacija, kiek metų apskritai buvo pragyventa kiekviename intervale, esant 100 000 naujagimių kohortai. Ši funkcija gali būti apskaičiuojama dviem atvejais: pirmuoju atveju turėtume žinoti mirties datą visų asmenų, mirusių šio amžiaus intervale. Toks skaičiavimų tikslumas svarbus medicinoje. Socialiniuose moksluose dažniausiai ši data nežinoma, tada manoma, kad mirčių skaičius išsidėsto tolygiai per amžiaus intervalus ir, pavyzdžiui, esant 5 metų intervalui, šiame intervale mirusieji asmenys išgyveno vidutiniškai 2,5 metų. Tada intervale pragyventų metų skaičių apskaičiuojame pagal formulę:

$${}_nL_x = n \cdot l_{x+n} + \frac{n}{2} \cdot {}_nd_x.$$

Pavyzdžiui, jeigu 30–35 metų amžiaus grupėje 96 819 asmenys sulaukė 35-ojo gimtadienio, tai jie kartu pragyveno $96819 \times 5 = 484095$ metus, o šio amžiaus intervale numirę 621 asmenys vidutiniškai kartu pragyveno $621 \times 2,5 = 1552,5$ metų. Visi kartu jie pragyveno $484095 + 1552,5 = 485647,5$ metų. Reikia neužmiršti, kad pirmieji intervalai trumpesni. Įvairių metodikų yra surašyta skaičiuojant paskutiniojo intervalo reikšmę. Šiuo atveju šį visą intervalą galima prilyginti 15 metų.

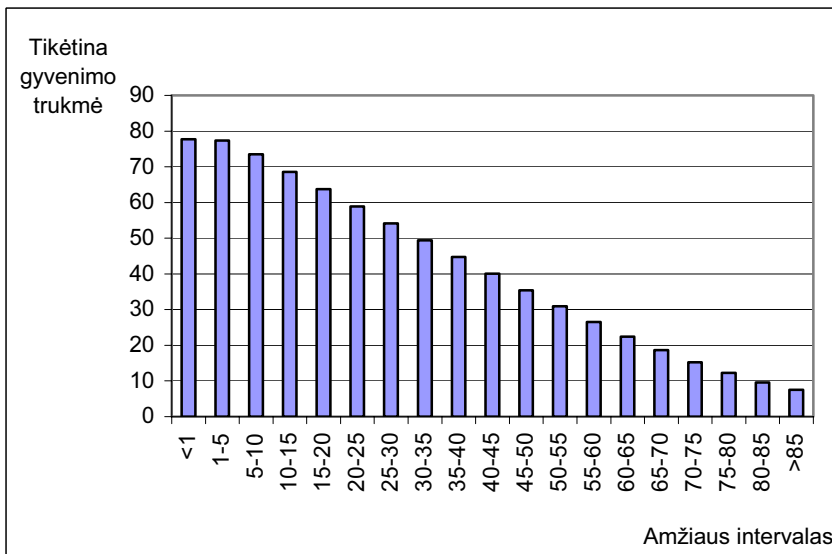
T_x : septintajame stulpelyje (${}_nT_x$) apskaičiuojama, kiek metų visai kohortai liko gyventi, kai sulauks x amžiaus. Šiuo tikslu ${}_nL_x$ reikšmės yra kaupiamos nuo apačios iki eilutės, atitinkančios x amžių. Taigi, T_x

stulpelis apskaičiuojamas pagal formulę: $T_x = \sum_{i=x}^w {}_nL_x$.

Pavyzdžiui, sulaukus 70 metų, kohorta dar turėtų gyventi $186855+225782,5+296245+353027,5=1061910$ metų.

e_x : aštuntajame (pagrindiniame) stulpelyje (e_x) pateikiama tikėtina gyvenimo trukmė gimstant, t. y. kiek metų vidutiniškai asmuo dar gyvens sulaukęs x gimtadienio. Pavyzdžiui, asmuo, sulaukęs 65-ojo gimtadienio, vidutiniškai dar gyvens 18,6 metų.

Apskaičiuotą tikėtiną gyvenimo trukmę galima pateikti stulpeline diagrama (3.9 pav.).



3.9 pav. Tikėtinų gyvenimo trukmės diagrama

Užduotys

3.8. 3.16 lentelėje pateikti duomenys apie 1985 m. Anglijos ir Velso moterų gyvenimo trukmės pasiskirstymą amžiaus intervaluose (šaltinis: *World Population Profile*, 1996). Sudarykite amžiaus lentelę ir apskaičiuokite tikėtiną gyvenimo trukmę.

3.16 lentelė. 1985 m. Anglijos ir Velso moterų gyvenimo trukmės pasiskirstymas amžiaus intervaluose

Metai	n	l_x	${}_n d_x$
<1	1	100000	825
1–5	4	99175	162
5–10	5	99013	89
10–15	5	98924	93
15–20	5	98831	139
20–25	5	98692	152
25–30	5	98540	179
30–35	5	98361	278
35–40	5	98083	432
40–45	5	97651	703
45–50	5	96948	1197
50–55	5	95751	1995
55–60	5	93756	3324
60–65	5	90432	5291
65–70	5	85141	7434
70–75	5	77707	10816
75–80	5	66891	14782
80–85	5	52109	18362
>85	10	33747	33747

3.9. Lietuvos statistikos departamento duomenys apie 2000 ir 2003 m. Lietuvos gyventojų amžių ir lytį pateikti 3.17 lentelėje. Nubraižykite 2000 ir 2003 m. amžiaus/lyties piramides ir apskaičiuokite darbingo amžiaus gyventojų demografinius krūvius. Gautus rezultatus palyginkite.

3.17 lentelė. Gyventojų skaičius pagal lytį ir amžių 2000 ir 2003 metų pradžioje

Amžius	Iš viso	Vyrai	Moterys	Iš viso	Vyrai	Moterys
	2003 m.	2003 m.	2003 m.	2000 m.	2000 m.	2000 m.
0	29813	15379	14434	36103	18591	17512
1–4	137159	70693	66466	153300	78743	74557
5–9	201554	103307	98247	250857	128637	122220
10–14	264439	135133	129306	291180	148257	142923
15–19	278778	141877	136901	270495	137721	132774
20–24	242033	122983	119050	258604	130030	128574
25–29	230042	115097	114945	281745	143896	137849
30–34	251389	124589	126800	279260	141961	137299
35–39	257463	127106	130357	302713	150250	152463
40–44	275645	132838	142807	263997	128196	135801
45–49	224937	106716	118221	229454	108319	121135
50–54	205044	93934	111110	192710	87890	104820
55–59	170918	75281	95637	203819	90384	113435
60–64	183654	77156	106498	190384	80027	110357
65–69	165572	65259	100313	174107	68755	105352
70–74	146131	53194	92937	141230	48178	93052
75–79	105892	31979	73913	86988	27387	59601
80+	92090	24783	67307	91575	26703	64872
Iš viso	3462553	1617304	1845249	3698521	1743925	1954596



4. STATISTINĖ GYVENIMO LYGIO ANALIZĖ: VISUOTINĖ GEROVĖ IR NELYGYBĖ

Pastaba. Dalį šio skyriaus apibrėžimų ir metodų galima rasti Pasaulio banko tinklalapyje

<http://www.worldbank.org/poverty/inequal/index.htm>

„Abu dalykai – pasitenkinimas gyvenimu ir jo trukmė – priklauso nuo socialinio teisingumo... Ilgą gyvenimo trukmę lemia ne absoliuti materialinė gerovė, bet tolygus gėrybių pasiskirstymas. Tai matyti ir pažvelgus į turtingas šalis. Švedijoje ir Japonijoje atotrūkis tarp didžiausių ir menkiausių pajamų yra mažiausias. Abiejose šalyse žmonės gyvena ilgiausiai, nors socialinė ir sveikatos sistemos labai skiriasi. Ir atvirkščiai, vis didėjanti socialinė nelygybė sietina su mažesne gyvenimo trukme...“

Stefan Klein. Laimės formulė

Gyventojų pajamų ir išlaidų lygis ir socialinis bei ekonominis gyventojų pasiskirstymas yra svarbiausi rodikliai, apibūdinantys gyvenimo lygį šalyje. Šiuolaikinėje visuomenėje gyventojų pajamų paskirstymas paklūsta ne tik griežtiems ekonomikos dėsniams. Jis yra socialinių procesų ir skirtingo lygio sutarčių tarp valstybės, verslininkų ir profsajungų rezultatas. Esant žemam pajamų lygiui ir palyginti didelei pajamų diferenciacijai daugėja žmonių, gaunančių labai mažas pajamas. Tokios situacijos padarinys – aukštas bendras skurdo lygis bei vaikus auginančių šeimų skurdas. Tai daro didelę neigiamą įtaką valstybės demografinėi situacijai, naujosios kartos ugdymui, ir yra viena iš pagrindinių nusikaltamumo didėjimą skatinančių priežasčių.

Visuotinės gerovės įvertinimui ir palyginimui naudojama socialinės gerovės funkcija, kurios idėją 1938 m. pasiūlė A. Bergson. Socialinės gerovės funkcija yra visuomenės materialinės gerovės matas, kuriuo vertinamas visų visuomenės narių ekonominis naudingumas:

$$W = f(x_1, x_2, \dots, x_n);$$

čia: W yra socialinė gerovė, x_i , $i = 1, 2, \dots, n$ yra kiekvieno visuomenės nario pajamos, n – populiacijos dydis.

Socialinės gerovės funkcija gali būti ne tik ekonominio naudingumo, bet ir kitokių kintamųjų, kurie siejasi su socialine gerove, pavyzdžiui, gyvenimo trukmės, funkcija. Socialinės gerovės funkcijos išraiška yra tokia, kad galėtų atspindėti visuomenės siekius. Paprasčiausia šios funkcijos forma yra $W = x_1 + x_2 + \dots + x_n$. Šiuo atveju maksimizuoti socialinės gerovės funkciją reiškia maksimizuoti visuomenės narių pajamas, neatsižvelgiant į tai, kaip šios pajamos visuomenėje yra pasiskirsčiusios. Net vieno visuomenės nario pajamų didėjimas, kai kitų visuomenės narių pajamos nesumažėja, socialinės gerovės funkciją didina.

Antra vertus, jei pritaikysime *maximin* naudingumo funkciją, gausime kitą socialinės gerovės funkcijos pavidalą:

$$W = \text{Max}(\text{minimum}(x_1, x_2, \dots, x_n)).$$

Šiuo atveju socialinė gerovė yra susijusi su neturtingiausių visuomenės narių pajamomis, ir gerovės maksimizavimas reiškia skurdžiausio visuomenės nario pajamų maksimizavimą, neatsižvelgiant į kitų pajamas. *Maximin* yra vienas iš sprendimo priėmimo kriterijų, kai sprendimą priimančias asmuo ieško su kiekvienu galimu veiksmu susijusios mažiausios galimos reikšmės (*minimumo*) ir nustato galimą didžiausią (*maksimalią*) šių minimalių dydžių reikšmę. Kadangi visuomenė, taip pat ir strategiją formuojanti vyriausybė paprastai teikia pirmenybę lygesniam pajamų pasiskirstymui prieš mažiau lygų, bet koks turtingesnių namų ūkių pajamų perdavimas skurdesniems ekonomikoje vertinamas kaip socialinės gerovės didinimas. Šis reiškinys vadinamas „perdavimo principu“ (*principle of transfers*).

Šios dvi socialinės gerovės funkcijos išreiškia du kraštinius požiūrius į tai, kaip visuomenė turi būti organizuojama, norint pasiekti didžiausią gerovę. Pirmoji socialinės gerovės funkcija pabrėžia bendras visuomenės pajamas, antroji – skurdžiausio visuomenės nario poreikius.

Su socialine gerove glaudžiai siejasi sąvokos *nelygybė* ir *skurdas*. Kartais skurdo indeksų apibrėžime įvedama nelygybė, taip pat nelygybė gali būti socialinės gerovės funkcijos argumentas. Nelygybės sąvoka yra daug siauresnė nei socialinės gerovės sąvoka. Nelygybė ignoruoja pajamų vidurkį ir išimtinai siejasi tik su pajamų sklaida (dispersija). Tačiau nelygybė savo ruožtu yra platesnė sąvoka nei skurdas. Nelygybė yra susijusi su visų individų pasiskirstymu, o skurdas suprantamas kaip kiekis asmenų ar namų ūkių, esančių žemiau tam tikros skurdo ribos.

Atliekant tarptautinio masto tyrimus apie ekonominės raidos įtaką socialinei gyventojų padėčiai vartojami įvairūs terminai ir rodikliai. Reikia pripažinti, kad ne visi jie vienareikšmiškai apibrėžti. Netgi tokie visuotinai vartojami terminai, kaip *vidurinė klasė* ar *skurdo riba*, neturi vieno apibrėžimo. Pavyzdžiui, *vidurinė klasė* dažnai apibrėžiama kaip sluoksnis gyventojų, kurių pajamos patenka į intervalą nuo 33 procentų mažiau visų šalies gyventojų pajamų medianos ir iki 50 procentų daugiau nei mediana. Tokių apibendrinančių rodiklių vertinimas valstybėms su skirtingomis ekonominėmis sistemomis ir skirtingu ekonomikos lygiu yra ganėtinai sudėtingas ir komplikluotas.

4.1. Skurdo sąvoka ir rodikliai

Valstybėje neturi būti nei didelio kai kurių piliečių skurdo, nei, antra vertus, turto, nes skurdas ir turtas gimdo vienas kitą. Skurdo riba tegul bus kaina skirtingo sklypo, kurį privalo turėti kiekvienas. Priėmęs šį matą, įstatymų leidėjas leidžia įsigyti turtą, kurio vertė didesnė du, tris, keturis kartus... Jei kažkas įsigyja daugiau, jis privalo perteklių atiduoti valstybei.

Platonas. Įstatymai, 744e

Skurdo sąvoka yra daugiaprasmė, ji kinta vystantis visuomenei ir skirtingai suprantama įvairiose šalyse. Pastaraisiais metais su skurdu

siejama **socialinės atskirties** sąvoka. Ši sąvoka integruoja skurdo, socialinių teisių netekimo bei netikrumo dėl jų sąvokas. Socialinės atskirties sąvoka pirmą kartą buvo pavartota 1974 m. Prancūzijoje, o nuo devintojo dešimtmečio pabaigos Europos komisija (vėliau ES) ją pradėjo plačiai vartoti kalbėdama apie socialines problemas, susijusias su ilgalaikie bedarbiais, žemos kvalifikacijos darbininkais bei imigrantais. Socialinė atskirtis kaip reiškinytis turi sąlygas egzistuoti daugėjant benamių, landynių gyventojų, ilgalaikių bedarbių, kai darbo ir pajamų neturi migrantai bei kai kurios etninės mažumos, kai didėja darbo netekimo rizika darbo rinkos naujokams.

Socialinės atskirties sąvoka vertinga tuo, kad apibūdina ne tik esamą situaciją, bet ir vykstančius procesus. Kiekvienoje visuomenėje egzistuoja dalis gyventojų, kurie nedalyvauja (yra atskirti) ekonominiame ir socialiniame gyvenime, negali be kitų pagalbos naudotis socialinėmis teisėmis, kenčia dėl nepakankamos savigarbos ar ilgalaikio pažeminimo. Sumažėjus pajamų mažiau tam tikro lygio, neproporcingai pradeda plisti pasitraukimas arba išstūmimas iš aktyvios narystės visuomenėje, t. y. didėja socialinė atskirtis.

Skurdo sąvoka kaip lemiamas socialinės atskirties veiksnys taip pat yra daugiaprasmiškas. Ji kinta vystantis visuomenei ir skirtingai suprantama įvairiose šalyse. Prieš keletą dešimtmečių skurdas būdavo apibrėžiamas kaip ekonominių išteklių stygius būtiniams poreikiams tenkinti. Išsivysčiusių šalių skurstantis žmogus silpnai išsivysčiusioje šalyje nebūtų priiskirtas prie skurstančiųjų grupės.

Pastaruoju metu skurdo sąvoka apibrėžiama bendriau. Skurstančiais individualiais ar šeimomis laikomi tie, kurių materialiniai, kultūriniai ir socialiniai ištekliai nepakankami savo šalyje susiklosčiusiam minimaliam gyvenimo būdui tenkinti.

Kadangi skurdas apibūdinamas įvairiai, tai ir skurdui įvertinti naudojami įvairūs rodikliai, būtent:

- *minimalus vartotojo krepšelis*,
- *minimalus gyvenimo lygis (MGL)*,
- *skurdo ribos: absoliuti, santykinė, subjektyvi*,
- *skurdo lygis ir skurdo gylis*.

Visi šie rodikliai susiję tarpusavyje. Minimalaus vartotojo krepšelio pagrindu skaičiuojamas MGL, kuris savo ruožtu yra absoliučios skurdo ribos nustatymo pagrindas.

Minimalus vartotojo krepšelis glaudžiai susijęs su vartojimo lygiu ir struktūra. 1990 m. Lietuvoje sudaryto minimalaus vartojimo krepšelio vertė buvo 55,4 proc. tuometinio faktinio vidutinio šalies vartojimo lygio. Krepšelis buvo taikomas pagrindinėms asmeninių poreikių grupėms. Jame buvo 52 maisto produktai, 55 ne maisto prekės ir 18 paslaugų. Sudarant šį rinkinį, buvo atsižvelgta į tuometinio vartojimo lygį ir struktūrą. Jau nuo 1993 m. krepšelyje nedetalizuojamos pramoninės prekės ir paslaugos ir, atsižvelgiant į tai, kad faktinio vartojimo didžiąją dalį sudaro išlaidos maistui, buvo nustatytas išlaidų santykis maistui ir kitoms prekėms bei paslaugoms: 80 proc. – maistui ir 20 proc. – kitoms prekėms (paslaugoms). Minimalaus maisto produktų rinkinio vidutinė vertė nustatoma pagal maisto produktų įsigijimo kainas mieste ir kaime bei atsižvelgiant į miesto ir kaimo gyventojų lyginamąjį svorį.

Taikomasis minimalus gyvenimo lygis (MGL) – tai vyriausybės oficialiai tvirtinamas dydis, pagrįstas tik valstybės biudžeto galimybėmis. Minimalus gyvenimo lygis apibrėžiamas taip: „Minimalus gyvenimo lygis (MGL) – tai šeimos mėnesinių pajamų suma, tenkanti vienam žmogui per mėnesį ir garantuojanti visiems minimalų socialiai priimtina poreikių patenkinimo lygį, atitinkantį organizmo maisto poreikius pagal fiziologines normas, taip pat minimalius drabužių, avalynės, baldų, ūkinių, sanitarijos ir higienos reikmenų, buto, komunalinių, buitinių, transporto, ryšių, kultūros ir švietimo paslaugų poreikius“ (1990 m. rugsėjo 27 d. ir 1994 m. lapkričio 3 d. redakcijos Lietuvos Respublikos gyventojų pajamų garantijų įstatymo 1 straipsnio 3 pastraipa). Tai visiškai atitinka ir visuotinės žmogaus teisių deklaracijos nuostatą, kurioje teigiama, jog „kiekvienas žmogus turi teisę į pakankamą gyvenimo lygį, kuris garantuotų jo ir jo šeimos gerovę, ir ypač jo maistą, drabužius, būstą, medicininę priežiūrą ir būtiną socialinį aptarnavimą“.

Lietuvoje MGL buvo nustatytas 1990 m., jo pagrindas – minimalus vartojimo krepšelis, nustatytas taikant statistinį normatyvinį metodą. Padidėjus kainoms, 1991 m. taikomasis MGL buvo indeksuotas 4 kartus; tai visiškai atitiko infliacijos tempą (kainoms padidėjus 4,8 karto, taikomasis MGL padidėjo 5 kartus). Tačiau 1992 m. kainoms padidėjus 12,6 karto, taikomasis MGL padidėjo tik 3,2 karto. Jo reali vertė 1992 m. sumažėjo 75 procentais. Šiuo metu taikomasis MGL yra netekęs savo prasmės ir tapęs rodikliu, su kuriuo siejami socialinių išmokų (šeimos pašalpos, našlaičio stipendijos, nėštumo pašalpos besimokančioms moterims, laidojimo pašalpos ir pan.) dydžiai.

Nuo 1993 m. Lietuvoje pradėta taikyti *skaičiuojamąją MGL*. Jis, kaip ir taikomas *MGL*, grindžiamas minimalių vartojimo poreikių tenkinimu, tačiau jo skaičiavimo metodas yra pakoreguotas. Skaičiuojamojo *MGL* pagrindas yra minimalaus maisto produktų rinkinio vertė (šio rinkinio maisto produktų kiekis ir sudėtis nustatyta atsižvelgiant į suaugusiųjų maisto produktų vartojimo normų ir vaikų racionalios mitybos normų duomenis, pateiktus namų ūkių biudžetų tyrimuose). Skaičiuojamojo *MGL* oficialiai netvirtina vyriausybė – jį kiekvieną ketvirtį nustato Socialinės apsaugos ir darbo ministerijos specialistai, remdamiesi Statistikos departamento namų ūkių biudžeto tyrimų rezultatais. Skaičiuojamojo *MGL* maisto produktų vertė remiamasi nustatant absoliučią skurdo ribą, taip pat nustatant valstybės remiamų pajamų dydį, kuris siejamas su minimaliu maisto produktų poreikio tenkinimu.

1999 m. skaičiuojamasis *MGL* vidutiniškai per mėnesį buvo 187,5 lito. Tačiau ir šiuo *MGL* negali būti remiamasi socialinei skurdo ribai nustatyti, arba jį taikant būtina daryti dideles išlygas. Vienas iš esminių šio *MGL* trūkumų – iškreipta vartojimo (šeimoms išlaidų) struktūra, mažinanti jo dydį: vadovaujantis Pasaulio banko rekomendacijomis, iš skaičiuojamojo *MGL* net 60 procentų šeimų išlaidų skiriama maistui. Šis *MGL* daugiau orientuotas į *fiziologinį* minimumą ir į žemutinę skurdo ribą.

Skurdo ribos

Absoliuti skurdo riba – tai nustatytas minimalus pajamų ir išlaidų lygis būtinausioms reikmėms patenkinti. Lietuvoje kaip atitikmuo šiai skurdo ribai yra minimalus gyvenimo lygis (*MGL*). *Absoliutus* skurdo riba turėtų būti apskaičiuojama remiantis minimaliomis asmeninio vartojimo normomis. Pavyzdžiui, JAV skurdo riba apskaičiuojama vidutinės maistui skiriamas išlaidas dauginant iš trijų. Šią metodiką 1960 m. pasiūlė Mollie Orshansky. Skurdo riba indeksuojama pagal vartotojo kainų indekso kitimą (įvertinus infliaciją). Šis skurdo ribos vertinimo būdas dažnai kritikuojamas, nes išlaidos būstui ir komunalinėms paslaugoms didėja greičiau nei maistui, taigi išlaidos maistui šeimų biudžete dabar sudaro mažesnę dalį nei 1960-aisiais.

Pasaulio bankas kaip didelio ir vidutinio skurdo standartą siūlo taikyti vieno ir dviejų dolerių pajamas per dieną. Jeigu šalies vidutinės pajamos itin didelės, tai net skurdžiausiai gyvenančios šeimos yra virš šios skurdo ribos.

Santykinė skurdo riba siejama su pajamų ar išlaidų vidutiniais rodikliais (aritmetiniu vidurkiu arba mediana). Ši skurdo riba visiškai priklauso nuo pajamų (išlaidų) pasiskirstymo visuomenėje ir kinta tiek keičiantis gyventojų pasiskirstymui pagal gyvenimo lygį, tiek augant ar smunkant vidutiniam gyvenimo lygiui šalyje. Pagal santykinio skurdo kriterijų prie skurstančiųjų priskiriami tie visuomenės nariai, kurių asmeninės pajamos daug mažesnės nei vidutinės. Pavyzdžiui, santykinio skurdo riba dažnai laikomas dydis – 50 (kartais nuo 40 iki 60) procentų asmeninių pajamų ar faktinių vartojimo išlaidų vidurkio arba medianos. Lietuvos santykinė skurdo riba yra 40 proc. vidutinių mėnesinių vieno namų ūkio nario išlaidų vidurkio (dėl netolygaus pajamų pasiskirstymo 1992–1995 metais 40 proc. vidurkis beveik sutapo su 50 proc. mediana).

Gyventojų pajamų medianos atskirose valstybėse labai skiriasi. Aukščiausia mediana yra JAV, taigi JAV žemiau skurdo ribos gyvenantys žmonės gali būti turtingesni negu trečiojo pasaulio šalių vidurinio-sios klasės atstovai. Santykinio skurdo skaičiavimas turi prasmę tose valstybėse, kuriose asmeninės pajamos, palyginti su objektyviomis minimaliomis normomis, yra didelis. Kitais atvejais galima gauti kuriozinius rezultatus, visiškai neatspindinčius tikrosios situacijos.

Subjektyvi skurdo riba nustatoma klausiant respondentų, kiek pinigų jiems užtektų minimaliems poreikiams tenkinti. Dažniausiai ši skurdo riba yra daug aukštesnė, negu apskaičiuota kitais metodais, kadangi subjektyvi skurdo riba nustatoma apklausiant asmenis, turinčius skirtingas pajamas. Turtingų gyventojų minimalūs poreikiai paprastai yra daug didesni nei skurstančiųjų.

Skurdo lygis – tai dalis asmenų, esančių žemiau skurdo ribos. Jų pajamos (vartojimo išlaidos) yra mažesnės nei tam tikru metodu apskaičiuota skurdo riba. **Skurdo gylis** – tai kiek vidutiniškai skurstančiųjų asmenų pajamos (išlaidos) yra mažesnės už skurdo ribą.

4.2. Gyventojų pajamų ir išlaidų tyrimai

Socialinių visuomenės teisingumo lygį atspindi ekonominis visuomenės narių susisluoksniavimas. Išsamiais tyrimais, atliekamais nuo XVIII amžiaus iki pat šių dienų įrodyta, kad šis socialinis bei ekonominis faktorius glaudžiai siejasi su kitais rodikliais ir ypač didelės įtakos turi šalies kriminogeninei situacijai. Neseniai atliktais tyrimais įrodyta,

kad pajamų tarp visuomenės narių pasiskirstymas susijęs su visuomenės narių sveikata. Didelis pajamų pasiskirstymo netolygumas mažina tikėtiną gyvenimo trukmę, didina mirtingumą ir net nutukimą. Kokie ryšiai sieja pajamų pasiskirstymo netolygumą su sveikatos rodikliais, iki šiol yra diskusijų objektas, tačiau įrodyta, kad pajamų diferenciacijos ryšys su amžiumi, rase, lytimi ir kitomis individualiomis socialinėmis bei ekonominėmis charakteristikomis yra tvirtas ir pastovus.

Gyventojų pajamų ir išlaidų dydis bei struktūra vertinama remiantis dviem nepriklausomais informacijos šaltiniais:

- gyventojų finansinių deklaracijų pajamų ir išlaidų balansu;
- namų ūkių biudžeto tyrimų rezultatais (arba panašaus tipo tyrimų rezultatais).

Namų ūkių biudžeto tyrimo objektas – namų ūkis, t. y. vienas asmuo ar asmenų grupė, kuri gyvena viename bute (name), turi bendrą biudžetą ir kartu maitinasi. Namų ūkiai atrenkami atsitiktinių imčių metodu iš gyventojų registro. Tokia atranka užtikrina vienodas galimybes visų visuomenės sluoksnių atstovams būti atrinktiems tyrimui.

Atrinkti namų ūkiai tyrime dalyvauja tik vieną mėnesį. Po mėnesio jie pakeičiami naujais. Tyrime dalyvaujantys namų ūkiai specialiose anketose surašo visas savo pajamas, išlaidas ne maisto prekėms ir paslaugoms apmokėti. Išlaidas maistui tiriamieji fiksuoja penkiolikai dienų (pusė namų ūkių – pirmąją mėnesio pusę, kita pusė – antrąją). Suvedant informaciją, išlaidos maistui perskaičiuojamos mėnesiui.

Dalis atrinktų namų ūkių tyrimuose nedalyvauja. Pavyzdžiui, 2001 metais namų ūkių biudžeto tyrimui buvo atrinkta 10560 namų ūkių, iš kurių tyrime dalyvavo 8001, t. y. 76 procentai. Penkiuose didžiausiuose miestuose nedalyvavo 36 procentai, kituose miestuose tyrime nedalyvavo 20 procentų, o kaime – 13 procentų atrinktų namų ūkių.

Pagrindinės sąvokos, vartojamos tiriant namų ūkių biudžetą

Namų ūkio tipas nustatomas pagal namų ūkio demografinę padėtį. Išskiriami šie tipai:

- vienišas asmuo;
- vienas suaugęs asmuo su vaikais iki 18 metų;
- sutuoktinių pora su vaikais iki 18 metų;
- kiti namų ūkiai su vaikais iki 18 m. Šiam namų ūkio tipui priklauso namų ūkiai, kuriuos sudaro tėvai su vaikais iki 18 m. ir vyres-

niais, kelių kartų namų ūkiai su vaikais iki 18 m., seneliai su vaikais iki 18 m. ir pan.

- sutuoktinių pora be vaikų;
- kiti namų ūkiai be vaikų.

Deciliai – (dešimtadaliai) skaičiuojami į dešimt lygių dalių padalijus eilutę, kurioje tiriami asmenys surašyti didėjimo tvarka pagal vartojimo išlaidų lygį. Pirmąjį dešimtadalį, t. y. pirmąjį decilį, sudaro namų ūkiai, kurių narių išlaidos yra mažiausios, antrąjį – namų ūkiai, kurių išlaidos yra didesnės nei pirmojo decilio tiriamųjų, bet mažesnės nei trečiojo ir t. t.

Disponuojamos pajamos – tai visos piniginės ir natūrinės pajamos, gautos už darbą iš ūkininkavimo, verslo, amatų, laisvosios prekybos, laisvosios profesinės veiklos, taip pat pensijos, įvairios pašalpos, stipendijos, pajamos iš turto, renta.

Vartojimo išlaidos – tai piniginės ir natūrinės išlaidos, skirtos namų ūkių vartojimo poreikiams tenkinti, t. y. išlaidos maistui, drabužiams, avalynei, būstui, sveikatos priežiūrai, kultūros, poilsio reikmėms.

Be gyventojų finansinių deklaracijų ir namų ūkių biudžeto tyrimų, privačios ir visuomeninės organizacijos atlieka įvairius rinkos ir gyventojų finansų tyrimus. Pateikiame fragmentą iš gyventojų finansų tyrimo, kurį atliko viešosios nuomonės ir rinkos tyrimų kompanija „TNS GALLUP“.

Anketa

Finansų monitoringas

1999 m. kovas

SIC
RINKOS TYRIMAI

ESANT GALIMYBEI, Į ANKETOS KLAUSIMUS ATSAKINĖJA ŠEIMOS NARYS, KURIO PAJAMOS SUDARO DIDŽIAJĄ ŠEIMOS PAJAMŲ DALĮ

1. Jūs gyvenate...?

atskirame name.....	1
atskirame bute.....	2
bute su bendra virtuve, bendrabutyje, kambaryje.....	3

2. Jūsų namas/butas, kuriame dabar gyvenate, yra Jūsų asmeninis turtas, ar Jūs jį nuomojate?

Asmeninis turtas: išpirktas už investicinius čekius	1
Asmeninis turtas: pastatytas jūsų pačių.....	2
Asmeninis turtas: gražintas valstybės, įgytas paveldėjimo būdu	3
Asmeninis turtas: pirktas	4
Nuomojamas iš valstybės ar vietos savivaldybės, iš įstaigos	5
Nuomojamas iš fizinio asmens.....	6

3. Kaip pasikeitė Jūsų šeimos finansinė padėtis, lyginant su praėjusių metų to paties laikotarpio padėtimi?

Jūsų šeimos finansinė padėtis...?

ženkliai pagerėjo	1
truputį pagerėjo.....	2
liko tokia pati.....	3
truputį pablogėjo.....	4
labai pablogėjo	5

4. Kaip Jūs manote, kokia bus Jūsų šeimos finansinė padėtis po 12 mėnesių?

Jūsų šeimos finansinė padėtis po 12 mėnesių...?

ženkliai pagerės	1
truputį pagerės	2
liks tokia pati	3
truputį pablogės	4
ženkliai pablogės	5

5. Kiek pinigų kiekvieną mėnesį lieka Jūsų šeimai, atėmus tokias būtinas išlaidas, kaip maisto, pirmojo būtinumo pramoninėms prekėms, išlaidas komunalinėms paslaugoms, elektrai, transporto išlaidas ir pan. *Atsakinėdami į pastarąjį klausimą turėkite galvoje, kad išlaidomis pirmojo būtinumo pramoninėms prekėms nelaikomos sąnaudų stambiais pirkiniais, statybinėms medžiagoms, įsiskolinimai, lizingo įmokos ir pan.*

Iki 50 litų	1
51 – 300 litų.....	2
301 – 500 litų.....	3
501 – 1000 litų.....	4

1001 – 2000 litų.....	5
Daugiau nei 2000 litų	6
Nelieka nė kiek.....	7

6. Kokiu tikslu, Jūsų nuomone, šiuo metu labiausiai vertėtų taupyti pinigus?

Taupyti kelis mėnesius, kad galima būtų įsigyti pirkinius, kurių negalėtumėte nusipirkti už vieno mėnesio atlyginimą	1
Taupyti 1– 2 metus, norint turėti lėšų stambioms išlaidoms.....	2
Taupyti keliolika metų, norint apsirūpinti ateičiai	3
Taupyti tik tuo tikslu, kad dėl visa ko nuolatos būtų pinigų	4
Šiuo metu Jūs nematote priežasties taupyti pinigus	5

7. Kokią pinigų sumą per mėnesį jūs galėtumėte atidėti taupymui?

Iki 50 litų	1
51 – 300 litų.....	2
301 – 500 litų.....	3
501 – 1000 litų.....	4
1001 – 2000 litų.....	5
Daugiau nei 2000 litų	6
Nė kiek	7

8. Koks taupymo būdas Jums atrodo pats tinkamiausias?

Turėti banke einamąją sąskaitą su ganėtinai žemomis metinėmis palūkanomis	1
Padėti pinigus į banką kaip kaupiamąjį indėlį su galimybe naudoti dalį pinigų savo poreikiams ir gauti didesnes palūkanas	2
Padėti pinigus į banką kaip terminuotą indėlį ir gauti aukštas metines palūkanas.....	3
Sudaryti draudimo sutartį (gyvybės, pensijinis draudimas)	4
Investuoti į vertybinius popierius	5
Investuoti į nekilnojamąjį turtą arba verslą (išskyrus vertybinius popierius).....	6
Laikyti grynuosius namuose.....	7
Kitas būdas (Prašome parašyti)	8

9. Ar Jūsų šeima turi kam nors skolinių įsipareigojimų (išskyrus pirkinius išsimokėtinai bei lizingo įmokas)?

Kokią sumą Jums dar reikia sumokėti?

Ne	1
Taip: iki 500 litų	2
501 – 2500 litų	3
2501 – 5000 litų	4
5001 – 10000 litų	5
10001 – 20000 litų	6
20001 – 50000 litų	7
daugiau nei 50000 litų	8

10. Ar Jūsų šeima moka už kokį nors pirkinį išsimokėtinai ar lizingo būdu?

Kokią sumą Jums dar reikia sumokėti?

Ne	1
Taip: iki 500 litų	2
501 – 2500 litų	3
2501 – 5000 litų	4
5001 – 10000 litų	5
10001 – 20000 litų	6
20001 – 50000 litų	7
daugiau nei 50000 litų	8

11. Ar Jūsų šeima skolino kam nors pinigų?

Jei taip, kokią sumą Jums dar turi gražinti?

Ne	1
Taip: iki 500 litų	2
501 – 2500 litų	3
2501 – 5000 litų	4
5001 – 10000 litų	5
10001 – 20000 litų	6
20001 – 50000 litų	7
daugiau nei 50000 litų	8

12. Kokios šiuo metu yra Jūsų šeimos santaupos grynaisiais arba indėliais banke? Čia reikia atmesti vertybinius popierius, pinigus, kuriuos Jums kas nors skolingas.

Įskaičiuokite pinigus, kuriuos Jums kas nors paskolino turite sumokėti už pirkinį išsimokėtinai.

Santaupų nėra	1
Santaupų yra: iki 500 litų	2
501 – 2500 litų	3
2501 – 5 000 litų	4
5 001 – 10 000 litų	5
10 001 – 20 000 litų	6
daugiau nei 20 000 litų.....	7

13. Ar Jūsų šeima turi vertybinių popierių (akcijų), kotiruojamų Nacionalinėje vertybinių popierių biržoje (Vilniaus banko, Hermio banko, AB „Lietuvos dujos“, AB „Lietuvos energija“, Biržų akcinės pieno bendrovės, AB „Kalnapilis“ ir pan.)

Taip.....	1
Ne	2

14. Ar Jūsų šeimos nariai yra sudarę pensijinio, kaupiamojo gyvybės draudimo sutartis?

Taip.....	1
Ne	2

15. Ar Jūsų šeima turi investicinių čekių?

Taip.....	1
Ne	2

16. Ar turite vertybinių popierių (akcijų) kurie nekotiruojami biržoje?

Taip.....	1
Ne	2

17. Ar žadate per artimiausius 12 mėnesių skolintis pinigų ar imti kreditą?

Taip, būtinai.....	1
Greičiausiai taip.....	2
Greičiausiai ne	3
Žinoma, ne	4
Sunku pasakyti.....	5

18. Jeigu per artimiausius metus skolinsitės pinigų (imsite kreditą), skolinsitės norėdami...

įsigyti nedidelį pirkinį	1
įsigyti ilgalaikio vartojimo prekių	2
įsigyti nekilnojamojo turto (pastatų, žemės)	3
remontuoti (rekonstruoti) gyvenamąją vietą	4
Kitu tikslu (prašome parašyti, koku)	5
Šiuo metu Jūs neketinate skolintis pinigų	6

19. Jeigu per artimiausius metus Jūs nuspręstumėte skolintis pinigų (imti kreditą), tai greičiausiai...

skolintumėtės iš draugų ar pažįstamų.....	1
paimtumėte paskolą iš banko	2
paimtumėte paskolą iš firmos, duodančios paskolas.....	3
vietoje kredito pasinaudotumėte lizingu.....	4

20. Kiek norėtumėte pasiskolinti?

Iki 5 000 litų.....	1
5 001 – 10 000 litų.....	2
10 001 – 25 000 litų.....	3
25 001 – 50 000 litų.....	4
daugiau nei 50 000 litų.....	5

4.3. Pajamų pasiskirstymo netolygumo įvertinimo metodai

Pajamų pasiskirstymo netolygumui įvertinti taikomi įvairūs metodai. Sukurta metodų, kuriuos taikant vertinamas pajamų pasiskirstymo netolygumas tam tikroje geografinėje srityje. Tačiau dažniausiai taikomi metodai pajamų pasiskirstymui tarp visuomenės narių įvertinti. Dažniausiai taikomi šie metodai: *Džini (Gini)* koeficientas, *struktūriniai* koeficientai, *Robino Hudo (Robin Hood)* indeksas, *Atkinsono (Atkinson)* indeksas, *Teilo (Theil) entropijos* indeksas.

Pasaulio banko duomenimis, pasaulyje egzistuoja didžiulis skirtumas tarp tų, kurie „turi“ ir tarp tų, kurie „neturi“, ir šis skirtumas per pastarąjį dešimtmetį itin padidėjo. Žemo išsivystymo lygio valstybėse

gyvena 40 procentų viso pasaulio gyventojų, tačiau šių šalių bendrasis vidaus produktas 1999 m. sudarė tik 11 procentų pasaulyje sukurto nacionalinio produkto. Ryškus kontrastas tam reiškiniui – aukšto išsivystymo valstybės, kuriose gyvena tik 15 procentų gyventojų, tačiau jos sukuria 56 procentų bendrojo pasaulinio nacionalinio produkto. Per pastarąjį dešimtmetį padidėjo pajamų diferenciacija ne tik tarp valstybių, bet ir valstybių viduje.

Struktūriniai koeficientai apskaičiuojami remiantis gyventojų pajamų (išlaidų) skirtumais. Pasaulinėje praktikoje dažnai taikomi deciliniai, kvartiliniai ir kvantiliniai diferenciacijos koeficientai, socialinio susisluoksniavimo koeficientas ir kt. Taikant decilinį koeficientą, gyventojų skaičius dalijamas į dešimt lygių dalių; taikant kvartilinį koeficientą, – į keturias dalis; taikant kvantilinį, – į penkias. Grupių pagal pajamas eilė tokia: pirmoji – su mažiausiomis pajamomis ir paskutinioji – su didžiausiomis pajamomis. Vienas iš dažniausių būdų diferenciacijai vertinti yra sukauptų (*kumuliatyviųjų*) dalių nustatymas pagal pajamas ir pagal gyventojų skaičių. Kartais pateikiama informacija, kurią dalį visų pajamų gauna 50 proc., 60 proc. ir 70 proc. gyventojų, gaunančių mažiausiai pajamų arba 1 proc., 5 proc. ir 10 proc. asmenų, gaunančių daugiausia pajamų. Pavyzdžiui, Pasaulio bankas 2002 m. paskelbė, kad 1 procentas turtingiausių pasaulio asmenų valdo tokio pat dydžio turta, kaip ir 57 procentai skurdžiausių pasaulio gyventojų.

Gini koeficientas yra dažniausiai naudojamas pajamų pasiskirstymo netolygumo matas. Gini koeficientas apskaičiuojamas iš Lorenzo kreivės. Ja vaizduojama, kokią sukauptų pajamų dalį gauna namų ūkiai, išrikiuoti eilėje nuo gaunančių mažiausias iki didžiausias pajamas. Jei pajamos pasiskirsčiusios tolygiai, Lorenzo kreivė tampa 45 laipsnių kampo įstrižaine – lygybės linija (4.1 pav.). Didėjant pajamų pasiskirstymo netolygumui, Lorenzo kreivė vis labiau tolsta nuo šios įstrižainės. Gini koeficientas yra santykis ploto, esančio tarp įstrižainės ir Lorenzo kreivės, bei ploto, esančio žemiau Lorenzo kreivės. Gini koeficientas kinta nuo nulio iki vieneto.

Pasaulio bankas pateikia pasirinktų 10 valstybių Gini koeficientus, kuriais remiantis vertinamas pajamų pasiskirstymo valstybių viduje netolygumas (4.1 lentelė).

Vertinant viso pasaulio gyventojų pajamų diferenciaciją Gini koeficientu, šis koeficientas skaičiuojamas taip, tarsi viso pasaulio gyventojai sudarytų vieną didelę tautą. Įvertinus 91 valstybės, kurių gyventojai su-

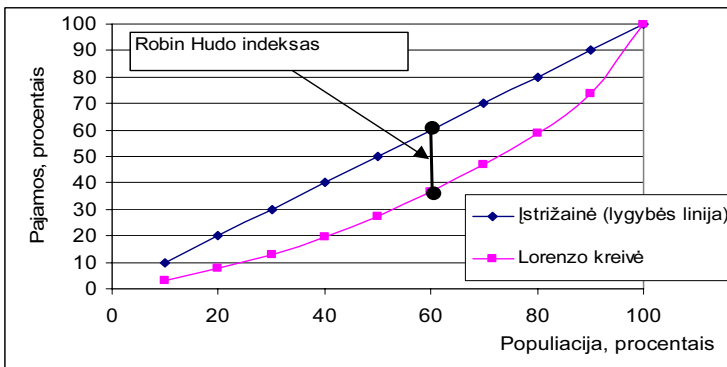
daro 85 proc. pasaulio populiacijos ir 95 proc. pasaulio pajamų, apskaičiuota globali Gini koeficiento reikšmė yra 66. Kaip suprasti tokią Gini koeficiento reikšmę? Ji yra analogiška situacijai, kai 66 proc. gyventojų pajamos lygios nuliui, ir 34 proc. pasidalina tarpusavyje visas gaunamas valstybėje pajamas.

4.1 lentelė. Kai kurių valstybių pajamų pasiskirstymo Gini indeksas

Valstybė	Tiriami metai	Gini indeksas
Brazilija	1997	59,1
Kanada	1994	31,5
Kinija	1998	40,3
Indija	1997	37,8
Japonija	1993	24,9
Meksika	1996	51,9
Rusijos federacija	1998	48,7
Pietų Afrika	1994	59,3
Jungtinė Karalystė	1991	36,1
Jungtinės Valstijos	1997	40,8

Šaltinis: *Worlds Bank: 2001 World Development Indicators*

Robin Hudo indeksas yra didžiausias vertikalusis atstumas tarp lygybės linijos ir Lorenzo kreivės (4.1 pav.). Šis indeksas apytiksliai parodo, kurią dalį visų pajamų namų ūkiai, gaunantys daugiau nei vidutinės pajamas, turėtų perduoti namų ūkiams, kurių pajamos mažesnės nei vidurkis, kad pajamos pasiskirstytų tolygiai.



4.1 pav. Lorenzo kreivė ir Robino Hudo indeksas.

Atkinsono indeksas yra vienas iš nedaugelio matų, kurie pateikia socialinės gerovės įvertinimą, paremtą socialinės gerovės funkcija. Indeksas įvertinamos vadinamosios „teisingos vidutinės pajamos“, kurios apibrėžiamos kaip toks pajamų vienam gyventojui lygis, kuris tenkintų kiekvieną gyventoją bei būtų išlaikomas bendras gerovės lygis toks, kokį užtikrina tikrosios bendrai sukurtos pajamos. Atkinsono indeksas A_ε diskretiniu atveju yra apskaičiuojamas pagal formulę:

$$A_\varepsilon = 1 - \left[\frac{1}{n} \sum_{i=1}^n \left[\frac{y_i}{\mu} \right]^{1-\varepsilon} \right]^{\frac{1}{(1-\varepsilon)}}, \quad (4.1)$$

čia: y_i yra visų pajamų dalis, kurią gauna i -toji grupė, μ – pajamų vidurkis, o ε yra vadinamasis nelygybės antipatijos parametras. Parametras ε įvertina visuomenės socialinio teisingumo suvokimą bei norą, kad pajamos būtų paskirstomos tolygiai. Šis parametras gali kisti nuo nulio iki begalybės. Kuo didesnė ε reikšmė, tuo visuomenė labiau susirūpinusi pajamų nelygybe. Jeigu $\varepsilon=0$, pajamų nelygybės nėra. Praktiškai ε reikšmė kinta intervale nuo 0,5 iki 2.

Kuo tolygiau pasiskirstę tikrosios pajamos, tuo y_i yra arčiau vidurkio μ , ir tuo mažesnė Atkinsono indekso A_ε reikšmė. Esant bet kokiam pajamų pasiskirstymui, šio indekso reikšmė kinta nuo 0 iki 1.

Teilo entropijos matas. Ši entropijos matą pasiūlė Theil (1967), ir jis išplaukia iš informacijos teorijoje naudojamos entropijos sąvokos. Entropijos matas T yra apskaičiuojamas taip:

$$T = \sum_{i=1}^n y_i [\log(y_i) - \log(1/n)], \quad (4.2)$$

čia: y_i yra i - tosios grupės pajamų dalis, n - grupių skaičius. Šis indeksas kinta nuo nulio iki begalybės. Kuo didesnė reikšmė (didesnė entropija), tuo tolygiau pasiskirsčiusios pajamos.

Pastaba. Nėra vieno „geriausio“ pajamų pasiskirstymo tolygumo mato. Šie matai yra susiję tarpusavyje (jų koreliacijos koeficientas didesnis nei 0,8). Skirtingi matai nevienodai reaguoja į pajamų perskirs-

tymą visuomenės sluoksniuose. Pavyzdžiui, Atkinsono indeksas labiau siejasi su skurdo paplitimu, Gini koeficientas mažai reaguoja į pajamų perskirstymą viduriniuosiuose sluoksniuose, Robin Hudo indeksas neįtaria pajamų perskirstymui, jei jis vyksta toje pačioje pajamų vidurkio pusėje ir pan. Metodą reikia parinkti pagal sprendžiamo uždavinio tipą ir formuluojamas hipotezes.

4.4. Pajamų pasiskirstymo netolygumo įvertinimo pavyzdys

Lietuvos statistikos departamento duomenimis, 1996 m. vidutinės mėnesinės išlaidos deciliuose buvo tokios (žr. 4.2 lentelę):

4.2 lentelė. Vidutinės mėnesinės išlaidos deciliuose 1996 m.

Decilio nr.	Vidutinės išlaidos deciliuose (Lt)	Gyventojų dalis
1	105,6	0,1
2	160,7	0,1
3	199	0,1
4	235,7	0,1
5	273,5	0,1
6	313,9	0,1
7	364,2	0,1
8	428,9	0,1
9	530,6	0,1
10	920,7	0,1

Užduotis. Apskaičiuokite decilinį ir kvartilinį diferenciacijos koeficientus, Lorenzo bei Gini koeficientus ir palyginkite gyventojų išlaidų struktūrą pagal minimalų gyvenimo lygį.

Metodiniai nurodymai

Decilinis diferenciacijos koeficientas yra kraštinių decilių santykis, kuris parodo, kiek kartų 10 procentų turtingiausių gyventojų pajamos (išlaidos) viršija 10 procentų skurdžiausių gyventojų pajamas (išlaidas):

$K_d = \frac{D_{10}}{D_1} = 920,7/105,6 = 8,72$. Taigi 1996 metais vidutinės 10 proc. tur-

tingiausių gyventojų pajamos viršijo 10 proc. skurdžiausiai gyvenančių gyventojų pajamas 8,72 karto.

Kvartiliai variacinę eilutę dalija į keturias maždaug lygias dalis, jie žymimi Q_1 , Q_2 , Q_3 . Antrasis kvartilis Q_2 sutampa su mediana bei dalija imtį į dvi dalis. Tuomet Q_1 yra apatinės dalies mediana, o Q_3 – viršutinės dalies mediana. Apytiksliai medianą Q_1 galima prilyginti trečio decilio vidutinei reikšmei, o Q_3 – aštunto decilio vidutinei reikšmei. Tada

kvartilinis diferenciacijos koeficientas: $K_{kv} = \frac{Q_3}{Q_1} = 428,9/199 = 2,155$.

Kvartilinis koeficientas apytiksliai atskleidžia vidurinėsios klasės diferenciaciją.

Šie koeficientai skaičiuojami tik iš dviejų decilių reikšmių, todėl jie neatspindi visų likusiųjų decilių. Korektiškesni yra rodikliai, apskaičiuoti iš pajamų (išlaidų) pasiskirstymo visuose deciliuose. Dažniausiai naudojami rodikliai yra Lorenzo kreivė, Lorenzo ir Gini koeficientai.

Lorenzo kreivė padeda įvertinti, kaip lygiai pasiskirsto pajamų, turto ar kitokių rodiklių reikšmės intervaluose. Ši kreivė yra sukauptų pajamų dalių (dažnių) grafinis vaizdas. Brėžiant Lorenzo kreivę tiriamoji seka išrikiuojama didėjančia tvarka $x_1 < x_2 < \dots < x_n$ ir apskaičiuojami

taškai $(h/n, L_h/L_n)$, čia $h=0, 1, \dots, n$, $L_0=0$ ir $L_h = \sum_{i=1}^h x_i$, n – intervalų skai-

čius. Jeigu visos sekos reikšmės yra vienodo dydžio, Lorenzo kreivė yra istrižainė, vadinama lygybės linija. Brėžiant Lorenzo kreivę veiksmai atliekami šia tvarka:

1. Kiekvienam kintamajam (rodikliui) apskaičiuojami santykiniai dažniai, t. y. sumos dalys.
2. Stebėjimai išdėstomi didėjančia tvarka pagal kintamąjį x .
3. Sudaroma sukauptųjų dažnių lentelė kiekvienam išrikiuotam kintamajam.
4. Nubrėžiama diagrama ir istrižainės atkarpa.

Lorenzo kreive parodoma, kurią pajamų dalį gauna tam tikra visuomenės dalis. Istrižainė parodo idealiai vienodą pajamų pasiskirstymą. Kuo toliau Lorenzo kreivė nukrypusi nuo istrižainės, tuo nelygiau pajamos pasiskirsčiusios. Šią kreivę galima transformuoti į Lorenzo koeficientą ir į Gini koeficientą. Lorenzo koeficientu įvertinami Lorenzo kreivės nuokrypiai; Lorenzo koeficientas apskaičiuojamas pagal formulę:

$$L = \frac{1}{2} \sum_{i=1}^n |x_i - y_i|. \quad (4.3)$$

Čia: x_i ir y_i – santykiniai dažniai, n – intervalų skaičius.

Gini koeficientas paprastai žymimas raide G ir yra lygus santykiui:

$$G = \frac{A}{A+B}. \quad (4.4)$$

Čia: A yra srities plotas tarp įstrižainės ir Lorenzo kreivės, o B yra srities plotas žemiau Lorenzo kreivės, taigi $A+B$ yra sritis, kurios riba – įstrižainė. Kai pajamos pasiskirsčiusios idealiai vienodai, srities A plotas (tarp įstrižainės ir Lorenzo kreivės) lygus nuliui ir Gini koeficientas lygus **nuliui**. Kai pajamos pasiskirsčiusios labai netolygiai, t. y. pirmųjų devynių decilių pajamos lygios nuliui, o visas pajamas gauna dešimtas decilis, srities B (žemiau Lorenzo kreivės) plotas yra nulis, Gini koeficientas – **vienetas**. Taigi Gini koeficientas kinta nuo 0 iki 1.

Kai dydžiai x išrikiuojami didėjančia tvarka, G gali būti apskaičiuojamas pagal formulę:

$$G = \frac{\sum_{i=1}^n (2 \cdot i - n - 1) \cdot x_i}{n^2 \cdot \mu}. \quad (4.5)$$

Čia: i – decilio eilės numeris, n – decilių skaičius ($n=10$), μ – vidutinių išlaidų deciliuose aritmetinis vidurkis.

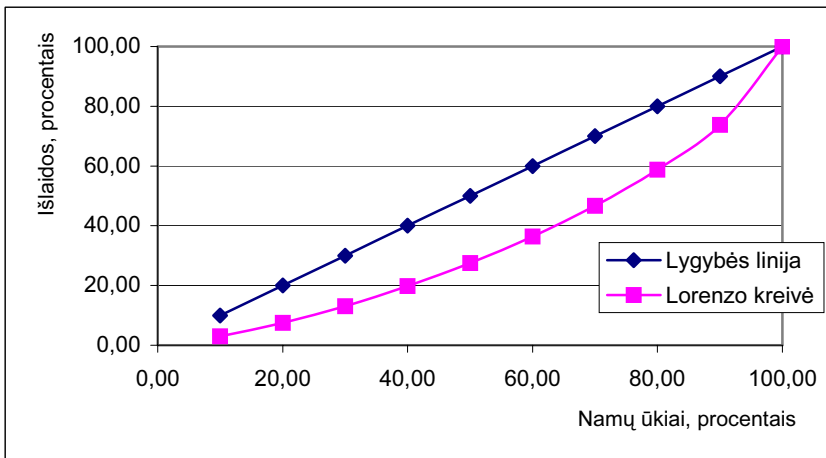
Reikia pabrėžti, kad Gini koeficientas neaprėpia visos informacijos apie pajamų pasiskirstymą. Dviem skirtingoms Lorenzo kreivėms apskaičiuotas Gini koeficientas gali būti vienodas. Pagrindinė šio koeficiento savybė yra ta, kad vieno kurio nors gyventojų sluoksnio pajamų didėjimas, kai tuo tarpu kitų sluoksnių pajamos nekinta, didina Gini koeficiento reikšmę.

Tarpiniai rodiklių skaičiavimo rezultatai pateikti 4.3 lentelėje. Brėždami Lorenzo kreivę (žr. 4.2 pav.), horizontaliojoje ašyje atidėsime sukauptuosius namų ūkius *cum F_p*, o vertikaliojoje ašyje – sukauptąsias išlaidas *cum F_i*.

4.3 lentelė. Tarpiniai skaičiavimo rezultatai

Decilio nr. i	Vidutinės išlaidos deciliose (Lt) x_i	Gyventojų dalis Fp_i	cum Fp	$D_i = x_i * Fp_i$	$F_{ti} = \frac{D_i}{\sum D_i}$	cum F_{ti}	$(2 \cdot i - 1) \cdot x_i$
1	105,6	0,1	0,1	10,56	0,030	0,030	-950,4
2	160,7	0,1	0,2	16,07	0,045	0,075	-1124,9
3	199	0,1	0,3	19,9	0,056	0,131	-995
4	235,7	0,1	0,4	23,57	0,067	0,198	-707,1
5	273,5	0,1	0,5	27,35	0,077	0,275	-273,5
6	313,9	0,1	0,6	31,39	0,089	0,364	313,9
7	364,2	0,1	0,7	36,42	0,103	0,467	1096,6
8	428,9	0,1	0,8	42,89	0,121	0,588	2144,5
9	530,6	0,1	0,9	53,06	0,150	0,738	3714,2
10	920,7	0,1	1,0	92,07	0,261	1,0	8286,3
Suma	$\sum x_i =$ 3532,8			$\sum D_i =$ 353,28	1,0		11500,6

Iš skaičiavimo rezultatų, grafiškai pateiktų 4.2 paveiksle matyti, kad tikrasis išlaidų pasiskirstymas labai skiriasi nuo idealaus. Didžiausi skirtumai yra kreivės viduryje, t. y. penktame, šeštame ir septintame deciliose.



4.2 pav. Lorenzo kreivė ir lygybės linija

Lorenzo koeficientą apskaičiuojame pagal (4.6) formulę:

$$L = \frac{\sum_{i=1}^n |Ft_i - Fp_i|}{2}, \quad (4.6)$$

čia: F_p – dalis gyventojų, kurių vidutinės išlaidos pateko į decilio intervalą;

F_t – dalis gyventojų, patekusių į decilį, kai išlaidos deciliuose apskaičiuotos pagal formulę:

$$Ft_i = \frac{x_i \cdot Fp_i}{\sum (x_i \cdot Fp_i)}. \quad (4.7)$$

Įrašę reikšmes iš 4.2 lentelės gauname Lorenzo koeficiento reikšmę:

$$L = (|0,030 - 0,1| + |0,045 - 0,1| + |0,056 - 0,1| + \dots + |0,261 - 0,1|) / 2 = 0,235.$$

Artimos nuliui Lorenzo koeficiento reikšmės parodo, kad gyventojų pajamos deciliuose tolygiai pasiskirsčiusios.

Gini koeficientas skaičiuojamas pagal 4.5 formulę. Nagrinėjamu atveju $n=10$. Apskaičiuavę reikšmių x_i aritmetinio vidurkio įvertį, gauname reikšmę 353,28. Įrašę reikšmes į formulę, gauname:

$$G = \frac{11500,6}{10 * 10 * 353,28} = 0,326 \text{ arba, padauginus iš šimto, apytiksliai } 33.$$

Išvados. Interpretuojant gautą Gini koeficiento reikšmę reikia atsiminti, kad ši apskaičiuota reikšmė yra analogiška situacijai, kai 33 proc. gyventojų pajamos lygios nuliui, ir 67 proc. gyventojų pasidalina visas gaunamas valstybėje pajamas tarpusavyje. Gautą reikšmę galime palyginti su kitų pasaulio valstybių Gini koeficiento reikšmėmis. Ryškūs Gini koeficiento pokyčiai per du pastaruosius dešimtmečius įvyko rytų Europos šalyse (4.4 lentelė)

4.4 lentelė. Kai kurių rytų Europos valstybių Gini koeficiento kitimas

Valstybė	Metai	Gini koef.	Metai	Gini koef.	Metai	Gini koef.
Bulgarija	1989	21,7	1993	33,1	1997	34,1
Čekijos respublika	1988	20	1992	23	1996	26
Vengrija	1987	24,4	1992	26	1998	25,3
Lenkija	1987	25	1993	29,8	1998	32,7
Rumunija	1989	23,3	1994	28,6	1997	30,5
Rusija	1991	26	1993	39,4	1996	45,5
Slovėnija	1987	19,5	1993	21,5	1996	26,3
Slovakija	1987	19,8	1993	24,1	1996	26,1

Šaltinis: Pasaulio bankas, 2003

Apibendrinus galima teigti, kad:

kai Gini koeficiento reikšmė mažesnė už 25, tai gyventojų pajamų diferenciacijos lygis yra labai žemas;

kai Gini koeficiento reikšmė didesnė už 25, bet mažesnė už 30, tai gyventojų pajamų diferenciacijos lygis yra nedidelis;

kai Gini koeficiento reikšmė didesnė už 30, tai jau esama esminių gyventojų pajamų diferenciacijos netolygumų;

Jungtinių tautų paskelbtoje Tūkstantmečio deklaracijoje šalys skirstomos į dvi grupes: pirmąją grupę sudaro valstybės, kurių Gini koeficientas neviršija 43, antrąją grupę, – kurių Gini koeficientas didesnis nei 43. Kai Gini koeficientas didesnis nei 43, šalyse yra labai daug skurstančių, todėl jose daug sunkiau sumažinti skurdą nei šalyse, kurių pajamų diferenciacija nėra tokia didelė. Šalių, kurių gyventojų pajamų diferenciacija didelė, ekonominio vystymosi greitis turi būti daug didesnis.

Užduotys

4.1. Lietuvos statistikos departamentas pateikia šiuos duomenis (žr. 4.5 lentelę) apie gyventojų vidutines mėnesines išlaidas deciliuose 1997, 1998, 1999 ir 2000 metais.

Nubrėžkite Lorenzo kreivę ir apskaičiuokite kiekvienų metų decilinius diferenciacijos koeficientus, Lorenzo ir Gini koeficientus. Palygin-

kite gyventojų išlaidų struktūros pagal pragyvenimo minimumą pokyčius. Remdamiesi gautais rezultatais parenkite 1997–2000 metų gyventojų pasiskirstymo pagal pajamas ir išlaidas analizę.

4.5 lentelė. Vidutinės mėnesinės išlaidos deciliuose 1997, 1998, 1999 ir 2000 metais

Deciliai	1997 m.	1998 m.	1999 m.	2000 m.
1	118,3	135,3	133,5	128,4
2	176,9	199,5	202,8	189,6
3	219,2	247,6	248,8	234,9
4	257,3	292,5	290,4	276,8
5	295,7	335,2	333,7	319,2
6	340,1	380,6	382,4	365,6
7	391,4	438,8	440,2	420,9
8	458,5	516,6	512,5	491,5
9	564,1	642,2	631,2	608,1
10	1004,6	1080,3	1077,7	1008,7

4.2. 4.6 lentelėje pateikti duomenys apie Rusijos gyventojų pasiskirstymą pagal vidutinės mėnesinės pajamas vienam asmeniui 1992 metais.

4.6 lentelė. Rusijos gyventojų pasiskirstymas pagal vidutinės mėnesinės pajamas vienam asmeniui 1992 metais

Vidutinės mėnesinės pajamos rubliais	Gyventojų dalis proc.
Iki 1000	4,7
1000,1–2000,0	21,8
2000,1–3000,0	23,0
3000,1–4000,0	17,0
4000,1–5000,0	13,5
5000,1–6000,0	6,5
6000,1–7000,0	4,3
7000,1–8000,0	2,9
8000,1–9000,0	1,9
9000,1–10000,0	1,3
Daugiau nei 1000	3,2

Apskaičiuokite decilinius diferenciacijos koeficientus, Lorenzo ir Gini koeficientus. Skaičiuodami naudokitės viduriniąja vidutinių mėnesinių pajamų intervalo reikšme. Atlikite Rusijos gyventojų išlaidų struktūros pagal pragyvenimo minimumą analizę, kai pragyvenimo minimumo reikšmė 980 rub.

4.3. 4.7 ir 4.8 lentelėse pateiktas Lietuvos gyventojų vidutinių mėnesinių disponuojamų (piniginių ir natūrinių) pajamų ir vidutinių mėnesinių vartojimo išlaidų (piniginių ir natūrinių) pasiskirstymas deciliuose 1999 metais.

Nubrėžkite pateiktų namų ūkio tipų Lorenzo kreivę ir apskaičiuokite kiekvieno namų ūkio decilinius diferenciacijos koeficientus, Lorenzo ir Gini koeficientus. Palyginkite gyventojų pajamų ir išlaidų struktūros pagal pragyvenimo minimumą skirtumus. Remdamiesi gautais rezultatais parenkite gyventojų pasiskirstymo pagal namų ūkio tipus analizę.

4.7 lentelė. Lietuvos gyventojų vidutinių mėnesinių disponuojamų (piniginių ir natūrinių) pajamų pasiskirstymas deciliuose vidutiniškai vienam namų ūkio nariui (Lt)

Namų ūkio tipas	Deciliai									
	I	II	III	IV	V	VI	VII	VIII	IX	X
Vienišas asmuo	150,6	293,8	341,8	383,7	428,6	481,8	552,5	648,5	820,8	1434,4
Vienas suaugęs su vaikais iki 18 m.	93,7	146,4	184,5	236,8	268,2	292,8	344,8	409,7	487,4	765,6
Poros su vaikais iki 18 m.	80,9	178,1	231,6	275,5	314,3	361,6	411,8	480,6	594,9	992,0
Kiti namų ūkiai su vaikais iki 18 m.	77,6	168,6	215,5	254,1	291,3	330,5	378,8	433,8	508,0	780,4
Poros be vaikų	149,8	289,3	343,1	368,7	430,9	478,8	546,1	647,1	804,8	1348,2
Kiti namų ūkiai be vaikų	131,3	240,0	292,0	340,6	388,2	436,9	495,4	567,7	693,9	1024,3
Lietuvoje	92,6	200,5	256,6	301,0	346,7	394,4	448,2	521,1	646,9	1072,2

4.8 lentelė. Lietuvos gyventojų vidutinių mėnesinių disponuojamų (piniginių ir natūrinių) išlaidų pasiskirstymas deciliuose vidutiniškai vienam namų ūkio nariui (Lt)

Namų ūkio tipas	Deciliai									
	I	II	III	IV	V	VI	VII	VIII	IX	X
Vienišas asmuo	204,2	292,5	352,1	405,9	457,7	517,0	598,8	694,2	855,3	1469,1
Vienas suaugęs su vaikais iki 18 m.	111,8	163,3	195,4	231,8	264,2	306,1	359,2	425,6	532,0	819,9
Poros su vaikais iki 18 m.	126,6	190,2	233,4	271,7	309,7	353,6	413,2	482,0	589,4	974,2
Kiti namų ūkiai su vaikais iki 18 m.	110,8	167,7	210,7	244,2	276,0	312,7	358,1	412,9	501,5	801,5
Poros be vaikų	196,9	278,7	327,9	371,5	419,4	463,8	521,2	610,0	756,2	1358,3
Kiti namų ūkiai be vaikų	152,1	222,6	267,1	309,5	351,9	397,0	452,7	522,7	634,5	1045,3
Lietuvoje	133,5	202,8	248,8	290,4	333,7	382,4	440,2	512,5	631,2	1077,7



5. UŽIMTUMO IR NEDARBO STATISTIKA

„Bedarbio laisvė miegoti rytais kiek tik širdis geidžia yra nieko verta, nes jis jos nepasirinko laisva valia. Jausdamiesi nereikalingi, žmonės kartu jaučiasi ir paniekinti. Jiems kyla grėsmė susirgti depresija ir visiškai apsileisti.“

Stefan Klein. Laimės formulė

Užimtumas – viena iš svarbiausių socialinių bei ekonominių rinkos ekonomikos problemų. Pagal Tarptautinės darbo organizacijos (TDO) standartus užimtumas ir nedarbas yra viena kitą papildančios charakteristikos. Ekonominės sistemos stabilumą užtikrina darbo paklausos ir pasiūlos balansas. Tačiau darbo paklausa dažniausiai viršija pasiūlą, ir tai lemia nedarbo atsiradimą. Ilgai trunkantis nedarbas yra didelė problema ir gyventojų skurdo šaltinis, sukeliantis ne vien ekonomines, bet ir galias psichologines problemas. Tyrimais yra įrodyta, kad ilgos trukmės nedarbas sukelia nevisavertiškumo kompleksą bei sumažintą savęs vertinimą. Po kurio laiko dingsta darbo įgūdžiai ir darbo motyvacija. Bedarbiui yra didesnė tikimybė susirgti psichikos ar kitomis streso sukeltomis ligomis, pavyzdžiui, širdies ir kraujagyslių, todėl tokie žmonės gyvena trumpiau. Nedarbas neigiamai veikia visus visuomenės narius – net ir tuos, kurie turi darbą. Tai galima paaiškinti pastarųjų baime darbą prarasti.

Kai nedarbo lygis yra aukštas, t. y. darbo paklausa labai viršija darbo pasiūlą, atsiranda grėsmė ekonominiam ir socialiniam visuomenės stabilumui. Todėl rinkos ekonomikos valstybėse užimtumas ir nedarbas nuolat tiriamas. Statistiniai duomenys būtini darbo rinkos reguliavimui, gyventojų socialinės apsaugos užtikrinimui, darbuotojų perkvalifikavimui.

Pagrindiniai užimtumo ir nedarbo statistikos uždaviniai yra šie:

- surinkti duomenis apie užimtųjų ir bedarbių skaičius;
- ištirti darbo rinkos padėtį ir nustatyti nedarbo lygį;

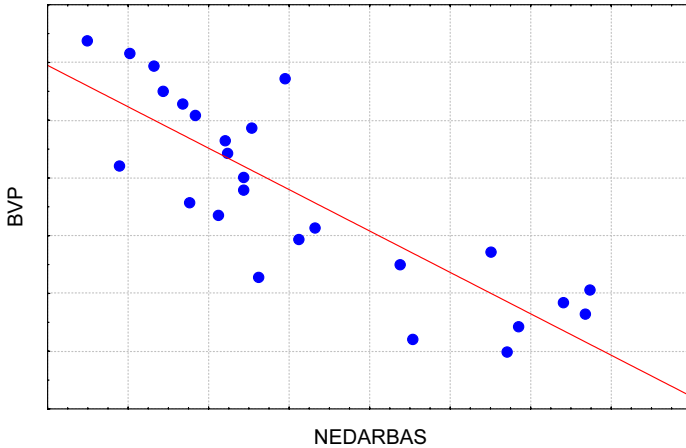
- ištirti gyventojų įsidarbinimo galimybes, kad būtų galima diegti gyventojų užimtumo programas;
- nustatyti ryšius tarp gyventojų užimtumo, pajamų, išsilavinimo ir motyvacijos gyventojų užimtumo programoms tobulinti.

Žvelgiant ekonomiškai, pagrindinė nedarbo kaina – nepagaminta produkcija. Bedarbiai nemoka mokesčių, taigi mažina valstybės biudžetą, o vyriausybė bedarbiams moka pašalpas. Be to, bedarbių perkamoji galia yra maža, ir tai neskatina ekonomikos augimo. Kai ekonomikos raida nesudaro galimybių sukurti tiek darbo vietų, kad patenkintų visus, kurie gali ir nori dirbti, prarandama tiek prekių ir paslaugų, kiek galėjo būti sukurta esant pakankamai darbo vietų. Ekonomistai šias prarastas prekes ir paslaugas apibrėžia kaip bendrojo vidaus produkto atsilikimą. Tai apimtis, kai faktiškas BVP mažesnis už potencialų BVP. Potencialus BVP nustatomas teigiant, kad yra natūralus nedarbo lygis esant normaliam ekonomikos augimo tempui.

Ekonomikos „konstantos“

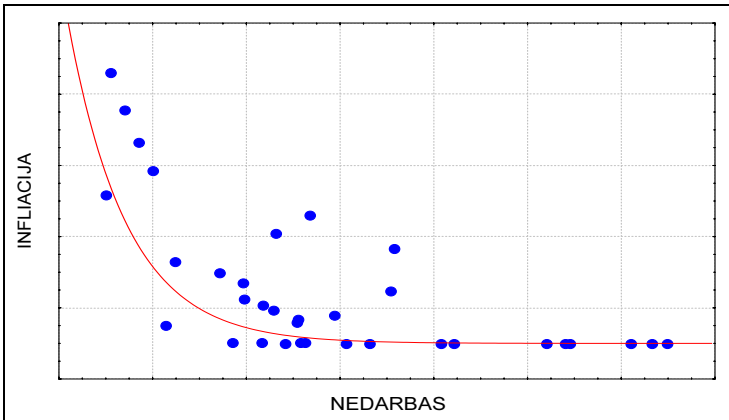
Pastaruoju metu ekonomistai, dažnai atsitiktinai, atranda empirinius ekonominio pasaulio dėsnius ir konstantas. Šių dėsnių patikimumas ir galiojimas priklauso nuo laikotarpio, per kurį dėsnis ar konstanta buvo stebima, ir nuo aplinkos įtakos. Vienas iš pavyzdžių, kurį pateikia ekonomikos pasaulis, yra vadinamasis *Okun* dėsnis. *Okun* nagrinėjo JAV ekonomikos duomenis. Tirdamas BVP kitimą nuo 1950 iki 1960m. *Okun* nustatė, kad dėl nepakankamos darbo jėgos paklausos, nedarbo lygiui padidėjus 1 procentu, BVP sumažėdavo 3 procentais. Grafiskai ryšį tarp BVP ir nedarbo lygio galima pavaizduoti diagrama (5.1 pav.). Tačiau šis dėsnis, kurį *Okun* pavadino „trijų dėsniu“, galiojo trumpai. Kylant darbo našumui ir išitraukus į gamybinę veiklą moterims, *Okun* konstanta nuo 1970 m. pradėjo kisti. Manoma, kad dabar ši konstanta apytiksliai yra 2,5.

Ekonomikos konstantos galioja tada, kai aplinkos sąlygos mažai kinta. Nuo 1950 iki 1960 m. *Okun* dėsnio konstanta nepakito nei nuo kintančių asmeninių sprendimų, nei nuo viešosios politikos, nei nuo įstatymų ir kultūrinių tradicijų. Tačiau po 1970 m., JAV prasidėjus svarbiems socialiniams ir ekonominiams pokyčiams, *Okun* dėsnio konstanta tapo nestabili.



5.1 pav. Okun dėsnis – ryšys tarp BVP ir nedarbo lygio

Kitą empirinę kreivę, apibūdinančią ryšį tarp bedarbystės ir infliacijos, atrado profesorius A.W. Phillipps. Analizuodamas empirinius duomenis nuo 1861 iki 1957 m. jis pastebėjo, kad nedarbas tiesiogiai siejasi su infliacija, ir ryšys apytikriai yra toks, koks pavaizduotas 5.2 pav. Tačiau 1970 m. buvo pastebėtas naujas reiškinys – infliacija ir nedarbas augo kartu. Šis reiškinys buvo pavadintas stagfliacija (t. y. sąstingis kartu su infliacija).



5.2 pav. Nedarbo ryšys su infliacija

Vienas iš nedaugelio ekonomikos dėsnių, išlaikiusių laiko išbandymą, yra Engel dėsnis. Šis dėsnis teigia, kad gaunant daugiau pajamų, mažėja pajamų dalis, išleidžiama maistui. Tačiau šis dėsnis nedetalizuoja pokyčių, ir todėl juo negalima remtis prognozuojant. Engel dėsnis naudingas tik bendram ekonomikos suvokimui. Pavyzdžiui, maisto ar medicinos pramonę gali dominti ne vien tai, kad augant pajamoms mažėja maistui išleidžiamų pajamų dalis, bet ir kaip keičiasi, tarkim, riebalų suvartojimas ar žmonių svoris. Todėl galima manyti, kad tvarus Engel dėsnio galiojimas siejasi su jo neapibrėžtumu ir paprastumu.

5.1. Bedarbiai ir nedarbo tipai

Kaip nustatyti, kiek gyventojų neturi darbo? Pagrindiniai metodai yra du:

1. Bedarbių registracija darbo biržose;
2. Darbo jėgos užimtumo ir nedarbo tyrimai (gyventojų apklausa).

Nuo 1980 iki 1990 m. (Lietuvoje iki 1994 m.) buvo taikomas pirmasis metodas – apskaičiuojamas darbo biržose užsiregistravusių asmenų skaičius. Šis metodas nėra tikslus, todėl pastaruoju metu dažniau taikomas antrasis metodas. Šiais metodais gaunami duomenys nesutampa (5.1 lentelė).

5.1 lentelė. Nedarbo lygis Lietuvoje 1998 – 2000 m.

Metai	Nedarbo lygis (darbo biržų duomenimis)	Nedarbo lygis (gyventojų apklausų duomenimis)
1998	6,4	12,2
1999	8,4	12,6
2000	11,5	13,3

Nagrinėjant nedarbą, išskiriama keletas jo tipų:

- ***Nepakankamos paklausos, arba ciklinis nedarbas.*** Šio tipo nedarbas atsiranda, kai stinga darbo jėgos paklausos ir negali įsidarbinti visi norintys ir galintys dirbti. Šis rodiklis priklauso nuo ekonomikos lygio. Augant ekonomikai, atsiranda darbo vietų, ir

šios rūšies nedarbas mažėja. Esant ekonomikos nuosmukiui, mažėja gamyba, taip pat mažėja ir darbo jėgos paklausa.

- **Sezoninis nedarbas.** Ryškiausias sezoninio nedarbo pavyzdys yra Kalėdų seneliai, kurie dirba tik labai trumpai. Sezoninis nedarbas būdingas tam tikroms gamybos ir paslaugų rūšims, būtent:
 - viešbučiams ir viešajam maitinimui;
 - turizmui;
 - žemės ūkio darbams (vaisių ir uogų rinkimui).
- **Paiėškos laiko nedarbas.** Kai žmogus netenka darbo arba jį keičia, tai trunka tam tikrą laiką. Šio tipo bedarbystė priklauso nuo darbo rinkos aktyvumo ir informacijos sklaidos efektyvumo.
- **Struktūrinis nedarbas.** Šio tipo nedarbas atsiranda keičiantis industrijos struktūrai. Vystantis pramonei, sukuriama naujos technologijos, o dalis senųjų tampa nebereikalingos. Šio tipo nedarbo paplitimas priklauso nuo keleto veiksnių:
 - darbo jėgos mobilumo;
 - ekonomikos kitimo greičio;
 - pramonės struktūros.

5.2. Nedarbas ir skurdas

Dažniausios skurdo priežastys yra ilgai trunkantis nedarbas ir nesugebėjimas gauti darbo su didesniu nei minimalus atlyginimu. Maži uždarbiai dažnai yra netinkamo arba nepakankamo išsimokslinimo pasekmė. Akivaizdu, kad nepakankamas išsilavinimas apriboja didesnio uždarbio galimybes ir šiai padėčiai ištaisyti vienintelis būdas yra geresnis bei tinkamesnis išsilavinimas. Tyrimais įrodyta, kad valstybės švietimo lygiui pakelti būtina sąlyga – decentralizavimas ir konkurencija. Svarbi sąlyga – individo suinteresuotumas ir galimybė pasirinkti išsilavinimo lygį ir apmokymo būdą. Pavyzdžiui, Vokietijos švietimo sistema ir stažuotių sistema labai padėjo sumažinti jaunimo nedarbo lygį.

Remiantis tyrimų rezultatais, galima teigti, kad valstybės remiamos vidutinio amžiaus asmenų mokymo programos retai kada duoda gerų rezultatų. Moterų dalyvavimas šiose programose šiek tiek didina užimtumą, tačiau įtaka nedarbui ir skurdui yra labai maža. Vyrų pasiekti rezultatai yra dar prastesni. Mokymo išlaidos dažniausiai yra didesnės nei gaunama nauda.

Problema yra ne tik išsilavinimas, bet ir darbo motyvacija. Nenoras dirbti taip pat yra viena iš skurdo priežasčių. Kartais dėl netinkamų socialinių programų kai kuriems visuomenės sluoksniams atsiranda galimybių išlaikyti jiems priimtina gyvenimo lygį nedirbant. Kartais socialinės išmokos didėja greičiau nei minimalus atlyginimas. Tokios pašalpos darbą daro nepatrauklų ir neskatina dirbti. Tokiam savanoriškam skurdui mažinti būtina tobulinti socialinio aprūpinimo programas.

5.3. Statistinės užimtumo ir nedarbo sąvokos ir rodikliai

Su gyventojų užimtumu susijusių sąvokų turinys šiek tiek skiriasi. Tai priklauso nuo užimtumo tyrimo metodikos.

Bedarbiai

Oficialiojoje statistikoje (darbo biržos duomenimis) bedarbiai yra nedirbantys darbingo amžiaus gyventojai, nesimokantys dieninėse mokyklose ir užsiregistravę valstybinėje darbo biržoje kaip ieškan-tys darbo ir pasirengę profesiniam mokymuisi.

Gyventojų užimtumo tyrimuose pagal EUROSTAT reikalavimus bedarbiai yra darbingo amžiaus gyventojai, kurie tiriamąją savaitę neturėjo darbo, o jį susiradę buvo pasirengę per artimiausias dvi savaites pradėti dirbti, keturias savaites intensyviai ieškojo mokamo darbo įvairiais būdais. Prie bedarbių grupės priskiriami ir tie moksleiviai, studentai, namų šeimininkės ir kiti neaktyvūs gyventojai, kurie tiriamąją savaitę aktyviai ieškojo darbo ir buvo pasirengę per artimiausias dvi savaites pradėti dirbti.

Užimtieji gyventojai

Užimtieji gyventojai yra tie, kurie tiriamąją savaitę dirbo kokį nors darbą ne mažiau kaip 1 valandą ir už jį gavo darbo užmokestį (pinigais arba maisto produktais bei kitais gaminiais) arba turėjo pelno (pajamų). Užimtieji kaime yra darbingo amžiaus žemdirbiai (turintys iki 3 ha žemės), kurie neturi kito pragyvenimo šaltinio, išskyrus tą žemę, nesvarbu ar joje užaugintą produkciją naudoja savo reikmėms, ar dalį jos parduo-

da. Smulkieji žemdirbiai – pensininkai prie užimtųjų kategorijos priskiriami tik tada, kai jų pajamos iš žemės ūkio yra didesnės nei gaunama pensija. Kiti pensininkai priskiriami prie neaktyviųjų gyventojų.

Neaktyvieji gyventojai

Gyventojai, nepriskirti prie užimtųjų ar bedarbių grupės, yra neaktyvieji gyventojai. Tai moksleiviai, studentai, namų šeimininkės, pensininkai, invalidai, nuteistieji, netekę vilties rasti darbo.

Darbo jėga (ekonomiškai aktyvūs gyventojai): visi užimtieji gyventojai ir bedarbiai.

Pagrindiniai statistiniai užimtumo ir nedarbo rodikliai yra šie:

- darbo jėgos aktyvumo lygis (gyventojų ekonominio aktyvumo koeficientas):

$$DJAL_1 = \frac{\text{Darbo jėga}}{\text{Darbingo amžiaus ir vyresni gyventojai}} \quad (5.1)$$

arba

$$DJAL_2 = \frac{\text{Darbo jėga}}{\text{Bendras gyventojų skaičius}} ; \quad (5.2)$$

- gyventojų užimtumo koeficientas:

$$GUK = \frac{\text{Užimtų gyventojų skaičius}}{\text{Darbingo amžiaus ir vyresni gyventojai}} ; \quad (5.3)$$

- nedarbo lygis (nedarbo koeficientas):

$$NL = \frac{\text{Bedarbių skaičius}}{\text{Darbo jėga}} ; \quad (5.4)$$

- bedarbių skaičius, tenkantis tūkstančiui užimtųjų:

$$BS_1 = \frac{\text{Bedarbių skaičius}}{\text{Užimtųjų skaičius}} \cdot 1000 ; \quad (5.5)$$

- vidutinė nedarbo trukmė (bedarbių skaičiaus pasiskirstymas darbo neturėjimo laiko intervalais):

$$T = \frac{\sum_i T_i N_i}{\sum_i N_i} . \quad (5.6)$$

Čia: T_i – darbo neturėjimo laikas i - toje grupėje, N_i – bedarbių skaičius i - toje grupėje.

Struktūrų skirtumus apibūdinantys koeficientai

1. Skirtumų indeksas:

$$I_{skirt} = \frac{1}{2} \sum |w_1 - w_0| . \quad (5.7)$$

Čia w_1 ir w_0 – rodiklių dalys (svoriai), kurie kinta nuo 0 iki 1. Indeksas I_{skirt} kitimo ribos irgi nuo 0 iki 1. Jei w išreikšta procentais, tai ši požymio dalis kinta $0 \leq w \leq 100$. Skirtumų indeksas šiuo atveju:

$$I_{skirt} = \frac{1}{2} \frac{\sum |w_1 - w_0|}{100} . \quad (5.8)$$

2. Absoliučių struktūrinių pokyčių tiesinis koeficientas:

$$d = \frac{1}{n} \sum |w_1 - w_0| . \quad (5.9)$$

Čia: n – grupių skaičius.

3. Absoliučių struktūrinių pokyčių kvadratinio nuokrypio koeficientas:

$$s = \sqrt{\frac{1}{n} \sum (w_1 - w_0)^2} . \quad (5.10)$$

5.4. Pagrindinių užimtumo ir nedarbo rodiklių skaičiavimo pavyzdys

Lietuvos statistikos departamentas atliko gyventojų užimtumo tyrimus. 5.2 lentelėje pateikti šio tyrimo metu gauti duomenys.

Šių duomenų pagrindu apskaičiuokite pagrindinius užimtumo ir nedarbo rodiklius.

5.2 lentelė. Gyventojų užimtumas 2001 m. gegužės mėn.

Rodiklis	2001 gegužės mėn.
Bedarbiai, registruoti darbo biržoje (tūkst.)	220,3
Bedarbiai, statistikos departamento duomenimis (tūkst.)	292,9
Užimtieji gyventojai, iš viso (tūkst.)	1481,8
Užimtieji gyventojai 15–64 metų (tūkst.)	1451,2
Neaktyvieji gyventojai (tūkst.)	1211,1
Darbo jėga (tūkst.)	1774,7
Gyventojai, metų pradžioje, (iš viso, tūkst.)	3692,6
15 m. ir vyresni gyventojai (tūkst.)	2985,8
15 – 64 metų gyventojai (tūkst.)	2483,1

1. Darbo jėgos aktyvumo lygis:

$$a) DJAL_1 = \frac{1774,7}{2985,8} = 0,594 \text{ arba } 59,4 \text{ procento;}$$

$$b) DJAL_2 = \frac{1774,7}{3692,6} = 0,481 \text{ arba } 48,1 \text{ procento.}$$

2. Gyventojų užimtumo koeficientas:

$$GUK = \frac{1481,8}{2985,8} = 0,496 \text{ arba } 49,6 \text{ procento.}$$

3. Nedarbo lygis (nedarbo koeficientas):

$$NL = \frac{220,3}{1774,7} = 0,124 \text{ arba } 12,4 \text{ procento.}$$

4. Bedarbių skaičius, tenkantis tūkstančiui užimtųjų:

$$BS_I = \frac{220,3}{1481,8} \cdot 1000 = 148,67.$$

5.5. Struktūros skirtumų vertinimo pavyzdys

Užimtųjų pasiskirstymas Lietuvoje pagal išsimokslinimą 1997 ir 2000 metais apibūdinamas 5.3 lentelėje pateiktais duomenimis.

Apibūdinkite skirtumus išsimokslinimo struktūroje 1997 ir 2000 metais, panaudodami šiuos rodiklius: skirtumų indeksą, absoliučią struktūrinių pokyčių koeficientą, absoliučią struktūrinių pokyčių kvadratinio nuokrypio koeficientą.

5.3 lentelė. Užimtųjų ekonomikoje gyventojų išsimokslinimas (tūkstančiais)

Grupės Nr.	Užimtųjų gyventojų išsimokslinimas (tūkst.)	1997	proc.	2000	proc.
	<i>Iš viso</i>	1570,7	100	1518	100
1	Aukštasis	314,3	20,0	316,7	20,9
2	Aukštesnysis	402,1	25,6	379,3	25,0
3	Vidurinis be prof. mok.	375,8	23,9	303,7	20,0
4	Vidurinis su prof. mok.	194,3	12,4	268,7	17,7
5	Devynmetis be prof. mok.	149,1	9,5	116,1	7,6
6	Devynmetis su prof. mok.	62,7	4,0	76,1	5,0
7	Pradinis, neturi pradinio	72,4	4,6	57,4	3,8

Tarpinio skaičiavimo rezultatai pateikti 5.4 lentelėje. Čia w_0 – tam tikro išsilavinimo lygio užimtųjų gyventojų skaičius 1997 metais (procentais), w_1 – tas pats 2000 metais (procentais).

Skirtumų indeksas $I_{skirt} = \frac{1}{2} \cdot \frac{14,41}{100} = 0,072$. Gauta indekso

reikšmė parodo, kad nuo 1997 iki 2000 metų skirtumai užimtųjų išsilavinimo struktūroje pakito neženkliai – apie 7 procentus.

5.4 lentelė. Tarpinio skaičiavimo rezultatai

Grupės Nr.	$w_0, \%$	$w_1, \%$	$ w_1 - w_0 $	$(w_1 - w_0)^2$
1	20,0	20,85	0,85	0,73
2	25,6	24,99	0,61	0,38
3	23,92	20,0	3,92	15,36
4	12,37	17,7	5,33	28,42
5	9,49	7,64	1,85	3,40
6	3,99	5,0	1,01	1,04
7	4,6	3,78	0,82	0,69
Σ	100	100	14,41	50,01

Absoliučių struktūrinių pokyčių tiesinis koeficientas $d = \frac{1}{n} \sum |w_1 - w_0| = \frac{1}{7} \cdot 14,41 = 2,05$. Vadinasi, vidutiniškai struktūriniai skirtumai yra 2,05 procento.

Absoliučių struktūrinių pokyčių kvadratinio nuokrypio koeficientas $s = \sqrt{\frac{1}{n} \sum (w_1 - w_0)^2} = \sqrt{\frac{1}{7} \cdot 50,01} = 2,67$. Šis rodiklis tiksliausiai parodo dirbančių asmenų išsilavinimo lygio skirtumą 1997 ir 2000 metais. Šis skirtumas yra 2,67 procento.

5.6. Gyventojų užimtumo, nedarbo bei kitų socialinių rodiklių ryšių analizė

Prognozuojant socialinius rodiklius, dažnai vertinami ryšiai tarp tiriamų požymių. Pavyzdžiui, gali būti tiriama, ar Lietuvoje dirbančių vyrų ir moterų atliekamas darbo pobūdis yra vienodas. Apklausus 1430 atsitiktinai parinktų vyrų ir moterų, kokio pobūdžio darbą jie dirba, gauti rezultatai pateikti 5.5 lentelėje.

5.5 lentelė. Dirbantys Lietuvos gyventojai

Lytis	Darbdaviai ir sau dirbantys	Samdomieji asmenys	Iš viso
Vyrai	147	562	709
Moterys	89	632	721
Iš viso	236	1194	1430

Remdamiesi šiais duomenimis nustatykite, ar vyrai ir moterys tiek darbdaviais, tiek samdomaisiais darbuotojais dirbo vienodai, t. y. ar užimtumas priklauso nuo lyties.

Tyrimo metodika

Kintamųjų, išmatuotų nominalioje skalėje arba rangų skalėje su nedideliu kategorijų skaičiumi, statistiniai ryšiai

Praktiškai atliekant bet kokį tyrimą pirmasis susipažinimas su duomenimis dažniausiai prasideda nuo dažnių lentelių. Pavyzdžiui, iš dažnių lentelės matyti, kiek tyrime dalyvavo vyrų, kiek moterų, kokia buvo respondentų šeiminė padėtis, koks išsilavinimas. Pirmasis žingsnis visada turėtų būti dažnių lentelių skaičiavimas, ypač kai analizuojami nominalieji ir ranginiai kintamieji.

Dažnių lentelės naudingos ir tada, kai reikia nustatyti esant ryšį tarp dviejų ar daugiau kintamųjų. Kartais požymiai būna susiję labai stipriai: jei profesija lakūnas, tai lytis paprastai vyras. Taigi šiuo atveju žinant požymį „profesija“ galima prognozuoti požymio „lytis“ reikšmę. Kitas kraštutinis – tai priklausomybės tarp požymių nebuvimas (jei akys žydros, tai kokia lytis?)

Ryšius tarp kintamųjų, išmatuotų *nominalioje* skalėje arba kintamųjų, išmatuotų *rangų* skalėje su nedideliu skaičiumi kategorijų, geriausia pateikti dažnių lentele. Sudarius dviejų ar daugiau kintamųjų dažnių lenteles, naudojami suderinamumo kriterijai (Pirsono χ^2 , Fišerio ir pan.), kuriais tikrinama, ar tarp stebimų ir tikėtinų dažnių yra reikšmingas skirtumas.

5.7. Pirsono χ^2 suderinamumo kriterijus

Išnagrinėkime vieną iš daugelio statistinių metodų, dažnai taikomų užimtumo ir nedarbo statistiniuose tyrimuose. Tyrėją domina, kaip tiksliai galima numatyti vieno požymio reikšmę, kai žinoma kito požymio reikšmė. Tačiau prieš sprendžiant šią problemą, reikia išspręsti kitą, paprastesnę: pirmiausia reikia nustatyti ar tarp šių požymių iš viso yra koks nors ryšys, ar jų kitimas nepriklauso vienas nuo kito? Tokį statistinio ryšio buvimą ar nebuvimą galima įvertinti specialiais matematiniais

metodais, iš kurių pats populiariausias – Pirsono χ^2 (Pearson chi-square) kriterijus. Šis kriterijus remiasi tuo, kad apskaičiuojami *tikėtini* dažniai, t.y. dažniai kurių reikėtų *tikėtis* tuo atveju, kai tarp požymių nėra jokio ryšio.

Pavyzdžiui, apklausta 20 vyrų ir 20 moterų, ar jie mieliau pasirinktų futbolo rungtynes, ar koncertą. Jei nėra jokio ryšio tarp lyties ir pomėgių, reikėtų *tikėtis* kad tiek futbolo rungtynes, tiek koncertą pasirinkusių vyrų ir moterų skaičius bus maždaug vienodas.

χ^2 kriterijaus taikymas susijęs su kai kuriais apribojimais, taikomais stebėjimų skaičiui n . Šio kriterijaus verčiau netaikyti, kai stebėjimų skaičius n mažesnis nei 30. Kita sąlyga, kad bent 75 proc. dažnių lentelės gardelių tikėtini dažniai būtų ne mažesni kaip 5. Jei šios sąlygos netenkinamos, rezultatai gali būti nepatikimi. χ^2 kriterijų galima taikyti ir 2×2 , ir didesnėms lentelėms.

χ^2 kriterijaus taikymą išnagrinėkime praktiškai. Tam pasinaudosi-me statistiniu tyrimu apie užimtųjų Lietuvos gyventojų, vyrų ir moterų, atliekamo darbo tipą, kurio rezultatai pateikti 5.5 lentelėje.

Reikia nustatyti, ar vyrai ir moterys vienodai renkasi darbo tipą, t.y. ar vienodai dirba tiek darbdaviais, tiek samdomaisiais.

Suformuluokime hipotezę. Hipotezė H_0 teigia, kad vyrai ir moterys vienodai dirba tiek darbdaviais, tiek samdomaisiais. Alternatyva H_1 – tarp darbdavių ir samdomųjų vyrų ir moterų skaičius skirtingas:

Hipotezė H_0 : pasirenkamo darbo tipas nepriklauso nuo lyties;

Hipotezė H_1 : pasirenkamo darbo tipas priklauso nuo lyties.

Statistinio reikšmingumo lygmens parinkimas. Tikrinant hipotezę dažniausiai taikomas toks principas: hipotezę atmetame, jei atsitinka tai, kas teisingos hipotezės atveju atsitikti praktiškai negalėjo. „Praktiškai negalėjo“ reiškia, kad tokio įvykio tikimybė yra labai maža. Tam įvedama reikšmingumo lygmens sąvoka. Matematikoje reikšmingumo lygmuo paprastai žymimas α raide, $\alpha > 0$. Sutariame įvykį laikyti praktiškai negalima, jei jo tikimybė mažesnė už α . Taigi tikrindami hipotezę darome eksperimentą ir hipotezę atmetame, jei įvykio tikimybė mažesnė už α . Tikimybė matematikoje apibrėžiama kaip dydis, kintantis nuo 0 iki 1. Taikomojoje statistikoje reikšmingumo lygmuo dažnai reiškimas pro-

centais. Kompiuterinėse duomenų analizės programose skaičiuojamas *mažiausias* reikšmingumo lygmuo, su kuriuo nagrinėjamu atveju teisinga nulinė hipotezė gali būti atmesta. Šis reikšmingumo lygmuo vadinamas *stebimuoju reikšmingumo lygmeniu* (*p-level*). Jeigu teigiama, kad skirtumai reikšmingi esant 5 proc. reikšmingumo lygmeniui (kai $\alpha \leq 0,05$) turima galvoje, jog egzistuoja 5 proc. tikimybė, kad rezultatai yra nepatikimi, t.y. esant teisingam rezultatui, daug kartų kartodami tyrimą, penkis kartus iš šimto galime suklysti ir teisingą rezultatą atmesti.

Jei klaidingo sprendimo tikimybė yra α , tai teisingo sprendimo tikimybė yra $1-\alpha$. Kuo mažesnė α reikšmė, tuo didesnė tikimybė, kad bus priimtas teisingas sprendimas.

Taigi galima pasakyti, kad reikšmingumo lygmuo – tai tikimybė atmesti hipotezę, kai ji yra teisinga. Tradiciškai tai vadinama pirmosios rūšies klaida. Antrosios rūšies klaida daroma tada, kai priimama hipotezė, nors iš tiesų ji yra klaidinga.

Reikšmingumo lygmuo, kuriuo remiantis hipotezė atmetama, yra sutartinis. Daugelyje mokslo sričių naudojamas 5 proc. reikšmingumo lygmuo. Tačiau reikia atminti, kad 5 proc. reikšmingumo lygmuo susijęs su ganėtinai didele klaidos tikimybe (5 proc.). Patikimesnis yra 1 proc. reikšmingumo lygmuo ($\alpha=0,01$), o 0,5 proc. ($\alpha=0,005$) ir 0,1 proc. ($\alpha=0,001$) reikšmingumo lygmenys dažnai laikomi labai patikimais. Apskritai, pagal priimtą terminologiją išvados, padarytos remiantis klaidos tikimybe $\alpha \leq 0,05$, laikomos reikšmingomis, remiantis klaidos tikimybe $\alpha \leq 0,01$, labai reikšmingomis, o esant klaidos tikimybei $\alpha \leq 0,001$, – maksimaliai reikšmingomis. Kartais literatūroje tokios išvados ženklinamos atitinkamai viena, dviem arba trimis žvaigždutėmis (5.6 lentelė).

5.6 lentelė. Klaidos tikimybės interpretacija

Klaidos tikimybė	Interpretacija	Ženklimas
$\alpha > 0,05$	Mažai arba visai nėra reikšmingų įrodymų prieš H_0	ns (<i>non-significant</i>)
$\alpha \leq 0,05$	Vidutiniškai reikšmingi įrodymai prieš H_0	*
$\alpha \leq 0,01$	Labai aiškūs įrodymai prieš H_0	**
$\alpha \leq 0,001$	Ypač reikšmingi įrodymai prieš H_0	***

χ^2 kriterijaus skaičiavimas

Analizuojant statistinius ryšius pagal χ^2 kriterijų, pirmiausia apskaičiuojami tikėtini dažniai, t.y., kokios turėtų būti lentelės gardelių reikšmės, jeigu tarp kintamųjų nebūtų statistinio ryšio. Du kintamieji yra tarpusavyje nepriklausomi, jei stebimi dažniai sutampa su tikėtinais dažniais. Tikėtini dažniai apskaičiuojami pagal formulę:

$$q_{ij} = \frac{f_{i*} f_{*j}}{n}; \quad (5.11)$$

čia q_{ij} - tikėtinas dažnumas gardelėje, esančioje i - tos eilutės ir j - tojo stulpelio sankirtoje,

$$f_{i*} - i\text{-tosios eilutės dažnių suma, } f_{i*} = f_{i1} + f_{i2},$$

$$f_{*j} - j\text{-tojo stulpelio dažnių suma, } f_{*j} = f_{1j} + f_{2j},$$

n - nagrinėjamų elementų skaičius, šiuo atveju $n=1430$.

Šioje formulėje naudojamų simbolių reikšmės paaiškintos 5.7 lentelėje.

Kadangi 5.5 lentelės pagrindinių (paryškintų) duomenų eilutės bei stulpeliai turi po dvi gardeles, tai tokia lentelė vadinama 2×2 , ir jos gardelės žymimos taip (5.7 lentelė):

5.7 lentelė. Simbolinis 5.5 lentelės reikšmių rašymas

Lytis	Darbdaviai ir dirbantieji sau	Samdomieji	Iš viso
Vyrai	f_{11}	f_{12}	f_{1*}
Moterys	f_{21}	f_{22}	f_{2*}
Iš viso	f_{*1}	f_{*2}	

Iš 2.4 lentelės duomenų apskaičiuojame tikėtinus dažnius:

$$q_{11} = \frac{709 \cdot 236}{1430} \approx 117, \quad q_{12} = \frac{709 \cdot 1194}{1430} \approx 592,$$

$$q_{21} = \frac{721 \cdot 236}{1430} \approx 119, \quad q_{22} = \frac{721 \cdot 1194}{1430} \approx 602.$$

Toliau apskaičiuojame skirtumus tarp išmatuotų reikšmių (5.5 lentelė) ir apskaičiuotų tikėtinų reikšmių. Jeigu hipotezė teisinga, šie skirtumai neturėtų būti dideli. Apskaičiuojamas χ^2 kriterijus:

$$\chi^2 = \sum_{i,j=1}^n \frac{(f_{ij} - q_{ij})^2}{q_{ij}}; \quad (5.12)$$

čia: f_{ij} – gardelės, esančios i -tosios eilutės ir j -tojo stulpelio sankirtoje, tikrasis (stebėtas) dažnis, kurio reikšmės nurodytos 5.5 lentelėje. Irašę reikšmes apskaičiuosime:

$$\chi^2 = \frac{(147 - 117)^2}{117} + \frac{(562 - 592)^2}{592} + \frac{(89 - 119)^2}{119} + \frac{(632 - 602)^2}{602} \approx 18,26.$$

Po to apskaičiuojami lentelės laisvės laipsniai:

$$df = (s-1)(t-1);$$

čia: df – lentelės laisvės laipsnių skaičius, s – eilučių skaičius, t – stulpelių skaičius.

Nagrinėjama lentelė sudaro dvi eilutes ir du stulpeliai. Taigi $df = (2-1)(2-1) = 1$.

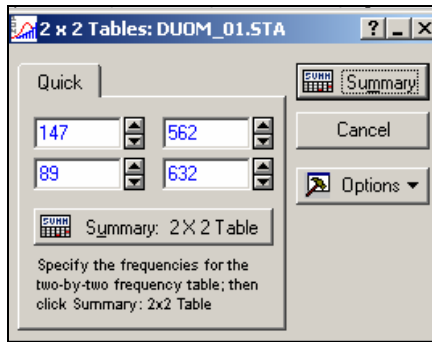
Statistinių išvadų formulavimas, skaičiavimui nenaudojant statistinės programos

Galiausiai reikia nustatyti, ar ši gauta kriterijaus χ^2 reikšmė yra pakankama, kad būtų galima teigti, jog tiriami požymiai yra priklausomi. Hipotezė apie kintamųjų nepriklausomumą yra atmetama, kai apskaičiuotos statistikos χ^2 reikšmė yra didesnė už kritinę reikšmę. Tarkime, pasirinkome reikšmingumo lygmenį $\alpha = 0,05$. Kaip apskaičiuoti kritinę reikšmę? Ją galima rasti matematinėse lentelėse, kurios pateikiamos prieduose (B lentelė, χ^2 skirstinio α lygmens kritinė reikšmė). Šiose lentelėse rasime, kad reikšmingumo lygmens $\alpha = 0,05$ kritinė reikšmė yra: $\chi_{0,05}^2(1) = 3,84$, ir ji yra daug mažesnė už gautą reikšmę 18,26, todėl hipotezę H_0 turime atmesti.

Nagrinėjama pavyzdyje kriterijaus reikšmė yra daug didesnė už kriterijaus reikšmę, atitinkančią reikšmingumo lygmenį 0,05 ar 0,005 ($\chi^2_{0,005}(1) = 7.88$). Vadinasi, atmetant hipotezę daroma labai menka klaida.

Skaičiuojant χ^2 kriterijų STATISTICA programa 2×2 lentelėms, paprasčiausias šios užduoties sprendimo būdas yra toks:

- meniu skyriuje *Statistics* pasirenkame punktą *Nonparametrics*;
- šio punkto analizės sąrašo pirmoji eilutė yra *2x2 Tables*. Pasirinkus šią eilutę, atidaromas duomenų įvedimo langas (5.3 pav.):



5.3 pav. 2×2 lentelės duomenų įvedimo langas

Į langą surašę turimas reikšmes, gausime išsamią rezultatų lentelę. Šios lentelės septintoje eilutėje (5.8 lentelė) yra kriterijaus χ^2 reikšmė (*Chi-square*) ir stebimojo reikšmingumo lygmens (*p*) reikšmė.

5.8 lentelė. χ^2 kriterijaus rezultatų lentelė.

	<i>Chi-square</i>	<i>df</i>	<i>p</i>
<i>Chi-square</i>	18,26	df=1	p=0,0000

5.8 lentelėje pateikta *Chi-square* – tai apskaičiuota χ^2 kriterijaus reikšmė, *df* – laisvės laipsniai, o *p* – stebimasis reikšmingumo lygmuo, t.y. kriterijaus reikšmę atitinkantis reikšmingumo lygmuo, pateiktas keturių skaitmenų po kablelio tikslumu.

Statistinių išvadų formulavimas, skaičiuojant statistinę programą

Skaičiavimui naudojant statistinę programą dėl hipotezės H_0 atmetimo sprendžiama pagal stebimąjį reikšmingumo lygmenį p . Šiuo atveju (5.8 lentelė) stebimasis reikšmingumo lygmuo p labai mažas ($p=0,0000$), todėl hipotezę, kad pasirenkamo darbo tipas nepriklauso nuo lyties, galima atmesti pasirinkus mažą reikšmingumo lygmenį (pavyzdžiui, $\alpha=0,005$).

Praktiškai, jei n labai didelis, daug didesnis už 1000, beveik visuomet hipotezė dėl požymių nepriklausomumo atmetama.

Jei tikėtini dažniai mažesni už 5, 2×2 lentelei galima taikyti *Fišerio* kriterijų.

5.8 χ^2 kriterijaus skaičiavimas didesnėms nei 2×2 lentelėms

Dažnai tenka nagrinėti požymius, kuriuos sudaro daugiau nei dvi kategorijos. Pavyzdžiui, norime sužinoti, ar yra priklausomybė tarp bedarbių lyties ir šeiminių padėties (5.9 lentelė).

5.9 lentelė. Duomenys apie bedarbių šeiminių padėčių

Šeiminė padėtis	Vyrai	Moterys	Iš viso
Susituokę	6	9	15
Nesusituokę	17	17	34
Išsituokę	7	5	12
Iš viso	30	31	61

Iki šiol tyrėme lenteles 2×2 , o ši lentelė yra 2×3 . Tyrimui galima naudoti χ^2 kriterijų, kuris apskaičiuojamas taip pat, kaip 2×2 atveju. Kuo didesnė lentelė, tuo sudėtingesni skaičiavimai, todėl jie atliekami pasitelkus kompiuterines programas. Išsiaiškinkime, kaip šią užduotį atlikti naudojant STATISTICA programą.

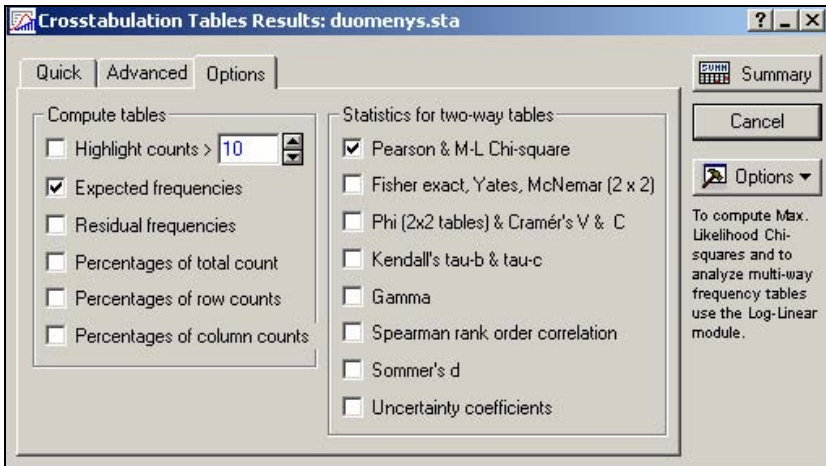
2×2 lentelių skaičiavimui naudojamas *Statistics* skyriaus *Nonparametrics* punktas, o didesnių nei 2×2 lentelių:

Statistics

Basic Statistics and Tables

Tables and banners.

Į pateiktą langą *negalima* įrašyti sugrupuotų duomenų, todėl naudojamos pradinės reikšmės, iš kurių sudaroma dažnių lentelė. Nurodžius kintamųjų pavadinimus, skaičiavimo parametrų parinkimo lange *Options* pažymima skaičiuoti tikėtinius dažnius ir χ^2 kriterijų (5.4 pav.).



5.4 pav. Dažnių lentelės parametrų parinkimo langas *Options*

Toliau (paspaudę *Summary* klavišą) jau gauname skaičiavimo rezultatus. Rezultatų lange pateikiama tikėtinų dažnių lentelė, apskaičiuota statistikos χ^2 reikšmė, laisvės laipsniai ir pirmos rūšies klaidos tikimybė (t.y., kokia klaida padaroma, atmetant nulinę hipotezę). Šiuo atveju rezultatai tokie:

$$\text{Pearson Chi-square: } 0,917; df=2; p=0,632$$

Primename, kad *Pearson Chi-square* – tai apskaičiuota χ^2 kriterijaus reikšmė, *df* – laisvės laipsniai, o *p* – kriterijaus reikšmė atitinkantis reikšmingumo lygmuo. Kadangi apskaičiuotas reikšmingumo lygmuo

($p=0,632$) didesnis už reikšmingumo lygmenį $\alpha=0,05$, tai hipotezės, kad bedarbių lytis ir šeiminė padėtis nesusiję, atmesti negalima.

5.9. Dažnių lentelių ryšio stiprumo matai

Φ koeficientas (*Phi coefficient*)

Apskaičiavus χ^2 kriterijų nustatoma, ar tarp požymių egzistuoja statistinis ryšys, tačiau nedaug galima pasakyti apie šio statistinio ryšio stiprumą. Norint tiksliau nustatyti ryšio stiprumą, naudojami kategorinių kintamųjų koreliacijos matai. Vienas iš tokių matų – Φ koeficientas. Šiuo matu įvertinamas koreliacinio ryšio stiprumas 2×2 lentelėse. Φ koeficiento reikšmės kinta nuo 0 (tarp požymių nėra jokio ryšio) iki 1 (tarp požymių yra labai stiprus ryšys).

Φ^2 koeficientas apskaičiuojamas χ^2 statistiką (5.12) padalinus iš viso duomenų skaičiaus n :

$$\Phi^2 = \frac{\chi^2}{n} \text{ arba } \Phi = \sqrt{\frac{\chi^2}{n}}. \quad (5.13)$$

Šis koeficientas gali būti naudojamas ir didesnėms nei 2×2 lentelėms skaičiuoti, tačiau jose kartais Φ koeficiento reikšmės būna didesnės už 1. Norint išvengti šio trūkumo, skaičiavimams naudojami papildomi koeficientai.

Kramerio V koeficientas (Cramer's V)

Kramerio V koeficientas apskaičiuojamas pagal formulę:

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(r-1, c-1)}}. \quad (5.14)$$

Čia: r – dažnių lentelės eilučių skaičius, c – stulpelių skaičius, $\min(r-1, c-1)$ parenkama mažiausia reikšmė. Šiuo kriterijumi įvertinama, kad eilučių ir stulpelių skaičius gali būti skirtingas. V koeficientas gali būti naudojamas didesnėms nei 2×2 lentelėms.

Kontingencijos koeficientas C (*Contingency coefficient*). Šis koeficientas dar vadinamas *Pirsono* kontingencijos koeficientu (*Pearson's contingency coefficient C*). Jam skaičiuoti taip pat naudojama χ^2 statistika:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}. \quad (5.15)$$

Kai kintamieji nepriklausomi, koeficientas C lygus nuliui. Tačiau šio koeficiento viršutinė riba priklauso nuo eilučių ir stulpelių skaičiaus. Pavyzdžiui, 2×2 lentelėms koeficiento C viršutinė riba yra 0,707. Didėjant eilučių ir stulpelių skaičiui, viršutinė riba irgi didėja, tačiau visada būna mažesnė už vienetą. Dėl šios priežasties C koeficientą sunkiau interpretuoti.

Užduotys

5.1. Užimtų ir bedarbių gyventojų pasiskirstymas pagal išsimokslinimą 2000 metais apibūdinamas šiais duomenimis (5.10 lentelė):

5.10 lentelė. Lietuvos gyventojų pasiskirstymas pagal išsimokslinimą

Išsimokslinimas	Užimtųjų (tūkst.)	Bedarbių (tūkst.)
Aukštasis	316,7	19,8
Aukštesnysis	379,3	59,9
Vidurinis be prof. mokymo	303,7	63,6
Vidurinis su prof. mokymu	268,7	61,3
Devynmetis be prof. mokymo	116,1	41,2
Devynmetis su prof. mokymu	76,1	24
Pradinis, neturi pradinio	57,4	5,9

Išsimokslinimo struktūroje apibūdinkite skirtumus tarp užimtųjų ir bedarbių.

5.2. Statistikos departamentas pateikė šiuos duomenis apie Lietuvos apskričių gyventojų užimtumą 1997 metais (žr. 5.11 lentelę).

5.11 lentelė. Lietuvos apskričių gyventojų užimtumas

Apskritis	Gyventojų	Bedarbių (tūkst.)	Nedarbo lygis (proc.)
Alytaus	202400	6,6	13,4
Kauno	757000	14,7	14,3
Klaipėdos	415400	9,7	13,3
Marijampolės	198500	5,1	7,2
Panevėžio	323300	9,5	12,6
Šiaulių	401900	12,9	13,3
Tauragės	130200	4,8	20,3
Telšių	182600	4,0	11,8
Utenos	202400	6,4	15,2
Vilniaus	895400	30,8	16,1

Pagal (5.1) – (5.6) formules apskaičiuokite:

- darbo jėgos dydį;
- darbo jėgos aktyvumo lygį;
- kiek bedarbių tenka 1000 užimtųjų;
- apibūdinkite ekonomiškai aktyvių gyventojų struktūrą.

5.3. 5.12 lentelėje pateikti duomenys apie nedarbą Rusijoje 2000 metų pradžioje:

5.12 lentelė. Duomenys apie nedarbą Rusijoje

Bendras gyventojų skaičius	147 (mln.)
Bendras bedarbių skaičius (tyrimų duomenys)	8,79 (mln.)
Nedarbo lygis	11,7 (proc.)
Oficialiai užregistruotų bedarbių	1 263 (tūkst.)
Gauna bedarbio pašalpą	1 090 (tūkst.)

Apskaičiuokite:

- darbo jėgos dydį;
- darbo jėgos aktyvumo lygį;
- kiek bedarbių tenka 1000 užimtųjų;
- kaip skiriasi oficialiai užregistruotų bedarbių skaičius nuo bedarbių skaičiaus pagal tyrimų duomenis;

- e) kiek procentų bedarbių negauna bedarbio pašalpos;
- f) apibūdinkite ekonomiškai aktyvių gyventojų struktūrą.

5.4. Gyventojų tyrimo metu apklausus 706 gyventojus buvo gauti šie duomenys (žr. 5.13 lentelę).

5.13 lentelė. Darbo ieškančių gyventojų pasiskirstymas pagal amžiaus grupes

Amžiaus grupė	Apklausti gyventojai	Ieško darbo
iki 20 m.	47	9
20–29	130	15
30–39	125	13
40–49	140	17
50–59	182	21
60 m. ir vyresni	82	6

Nustatykite, ar darbo ieškančių gyventojų skaičius priklauso nuo amžiaus grupės.



6. INTEGRUOTI SOCIALINĖS RAIDOS RODIKLIAI

Pastaba. Dalį šio skyriaus apibrėžimų ir skaičiavimo metodų galima rasti Jungtinių Tautų ataskaitos apie socialinę pažangą tinklalapyje http://hdr.undp.org/reports/global/2003/indicator/index_indicators.html

Integruoti arba daugybiniai rodikliai naudojami tarptautiniams palyginimams bei atskirų valstybių vidaus progresui vertinti, įvykių raidai analizuoti ir prognozuoti, turi keletą bendrų savybių. Šie rodikliai sudaro galimybę visumą sudarančias dalis (valstybes, provincijas ar sistemas) lyginti ir jų kitimą vertinti. Savo pobūdžiu integruoti rodikliai yra daugiamačiai, o jų skaičiavimo metodai yra susiję su taškų skaičiavimu ir rangų suteikimu.

Iki 1970 metų BVP buvo vyraujantis ekonominio progreso ir socialinės gerovės matas. Tačiau aštuntajame dešimtmetyje tapo akivaizdu, kad, vertinant gyvenimo lygį atskirose valstybėse, nepakanka įvertinti tik surinktą prekių ir paslaugų kiekį, kadangi pajamos yra tik priemonė socialinei pažangai pasiekti. Tai sudarė sąlygas dviejų tipų rodikliams, naudojamiems tarptautiniuose palyginimuose, atsirasti. Viena vertus, vertinant BVP vienam gyventojui, buvo pradėta skaičiuoti gyventojų perkamoji galia arba *realus BVP per capita*. Antra vertus, buvo sukurta ir papildomų matų progreso raidai vertinti. Viena iš pirmųjų gerovės vertinimo kriterijų 1979 m. sukūrė profesorius Morris ir pavadino jį fizinės gyvenimo kokybės indeksu (*Physical Quality of Life Index – PQLI*). Šiuo indeksu vertinami trys rodikliai: galimybė ilgai gyventi, vaikų mirtingumas ir raštingumas. Skaičiuojant *PQLI*, jį sudarantys rodikliai transformuojami į intervalą 0 – 1. Pavyzdžiui, didžiausias vaikų mirtingumas – 1, o mažiausias – 0, vadinasi, visos kitos reikšmės taip pat patenka į šį intervalą. Visos šių trijų rodiklių vienodo „svorio“ reikšmės sudedamos – tai ir yra šalies *PQLI* indeksas. Iš tikrųjų šiuo indeksu matuojama ne tiek socialinė raida, kiek gyventojų pagrindinių poreikių tenkinimas.

Iki pat 90-ųjų metų naudojamų integruotų rodiklių teorinio pagrindimo nebuvo. Devintajame dešimtmetyje buvo sukurti socialinių rodik-

lių naudojimo ir socialinės gerovės vertinimo teoriniai pagrindai. Šioje srityje didelis indų kilmės ekonomisto ir Nobelio ekonomikos premijos laureato Amartija Seno (*Amartya Sen*) darbo indėlis. Jo darbuose vystoma ir grindžiama idėja, kad vertinant socialinę gerovę svarbiausia yra ne tiek ekonominė nauda ar sukurta prabanga, kiek sukurta aplinka, skatinanti asmenybės gebėjimus ir raidą. Anot A. Seno teorijos, asmenybei funkcionuojant asmens įgytų vartojimo reikmenų vektorius yra konvertuojamas į šių reikmenų charakteristikų vektorių. Asmeninė daiktų „naudojimo funkcija“ iš charakteristikų vektoriaus sudaro funkcionavimo vektorių, kuris atspindi turimų vartojimo reikmenų naudojimo šabloną, t.y., kaip asmuo naudoja turimus daiktus. Funkcionalumas ir suteikiamos galimybės vertinamos atrinktais svarbiausiais rodikliais pagal tai, kaip jie atitinka tikslus. Pavyzdžiui, didelis valstybėje dirbančių gydytojų skaičius nereiškia, kad visi gyventojai naudojami gera medicinine priežiūra. Vienas iš pagrindinių rodiklių, glaudžiai susijusių su gera medicinine priežiūra, yra tikėtina gyvenimo trukmė. Taigi tikėtina gyvenimo trukmė tiesiogiai siejasi su socialine gerove.

6.1. Plačiausiai naudojami integruoti rodikliai

Šiuo metu plačiai naudojami Jungtinių Tautų sukurti integruoti socialiniai rodikliai. Vienas iš svarbiausių rodiklių yra socialinės pažangos indeksas (*Human Development Index – HDI*), kuris pirmąkart buvo pateiktas 1990 m. ataskaitoje apie socialinę pažangą (*Human Development Report – HDR*) JT raidos programoje (*United Nations Development Programme – UNDP*). Pagal šią programą kasmet pateikiama socialinės pažangos ataskaita, kurioje akcentuojami socialiniai rodikliai, priešingai nei Pasaulio banko ataskaitoje apie pasaulio raidą (*World Development Report – WDR*), kurioje dėmesys skiriamas ekonominiams rodikliams.

Be socialinės pažangos indekso *HDI*, dažnai skaičiuojami gyventojų skurdo indeksai (*HPI₁* ir *HPI₂*). Neseniai šie rodikliai buvo papildyti kitais integruotais rodikliais: raidos, įvertinant lyčių faktorių (*GDI* ir *GEM*) ir technologijų raidą (*Technology Achievement Index – TAI*).

Socialinės pažangos indeksas (*HDI*) yra glaudžiai susijęs su BVP vienam gyventojui, tačiau atskirose valstybėse jo reikšmės labai skiriasi. Tai aiškinama valstybių socialinės politikos skirtumais ir skirtingu pajamų pasiskirstymu.

Integruotais rodikliais vertinama valstybių ar regionų pažanga plačiai aptariama ir diskutuojama. Tai skatina vyriausybes ir administratorius sutelkti jėgas socialinės pažangos strategijų kūrimui ir įgyvendinimui bei žmonių gyvenamosios aplinkos gerinimui.

Integruotų rodiklių konstravimo schemas yra panašios. Šie rodikliai sudaromi iš kelių bazinių rodiklių:

- socialinės pažangos indeksas *HDI* yra tikėtinos gyvenimo trukmės, išsilavinimo įgijimo ir bendrojo vidaus produkto svartinė kompozicija.
- gyventojų skurdo indeksais (*Human Poverty Index – HPI₁* ir *HPI₂*) skurdas vertinamas keliais aspektais;
- lyčių lygybės raidos indeksą *GDI (Gender-related Development Index – GDI)*, kaip ir *HDI*, sudaro tikėtina gyvenimo trukmė, išsilavinimas ir bendrasis vidaus produktas, tačiau šie rodikliai skaičiuojami atskirai vyrams ir atskirai moterims.
- priešingai nei *GDI*, kitu indeksu lyčių lygybei vertinti *GEM (Gender empowerment measure – GEM)* vertinamas dalyvavimas administravime, profesiniame ir parlamentiniame darbe bei sprendimų priėmime.

Integruotų rodiklių konstravimo koncepcija ne visada būna tobula, o rangavimo interpretavimas dažnai lieka miglotas. Vienas iš esminių reikalavimų, keliamų integruotiems rodikliams – pagrįstumas. Pagrįstumas nagrinėjamas trimis aspektais: turinio pagrįstumas, konstravimo schemas pagrįstumas ir prognozavimo pagrįstumas. Integruotų rodiklių kaip daugialypių matų turinio pagrįstumą nulemia teorinės žinios, kurios siejasi su tiriamu reiškiniu. Konstravimo pagrįstumas turėtų būti patikrintas statistiniais metodais. Šie metodai taikomi tiriant bazinių rodiklių (komponentų) ryšius ir atrenkant didžiausios įtakos turinčius rodiklius. Atrinkti baziniai rodikliai siejami į visumą su parinktais svoriais. Teoriškai svoriai turi įkainoti funkcionalumo vektorių. Baziniai rodikliai gali būti siejami tiek tiesiniu, tiek ir netiesiniu būdu. Siejimo būdai ir komponentų svoriui nustatyti taikomi koreliacinės, regresinės ir faktoriinės analizės metodai. Sudarytų integruotų kriterijų prognozavimo objektyvumas tikrinamas gautus rezultatus lyginant su žinomomis reikšmėmis.

6.2. Socialinės pažangos indeksas

Socialinės pažangos *indeksas (HDI)* yra sudėtinis indeksas, kuris apibendrina tris pagrindinius gyventojų gyvenimo kokybės komponentus:

1. Galimybę ilgai gyventi, matuojamą tikėtina gyvenimo trukme gimstant (p_1).
2. Informacijos gyventojams prieinamumą (galimybę gauti mokyklinį išsilavinimą ir profesinį parengimą), kuri apibūdinama du rodikliai: suaugusių gyventojų raštingumas (p_2) ir dalis besimokančio jaunimo iš bendro jaunimo skaičiaus (p_3).
3. Ekonominės gyventojų galimybės, atsispindinčios bendrajame vidaus produkte (BVP) apskaičiuotame vienam gyventojui (*per capita*). Tai atskleidžia galimybę pasiekti tokį pajamų lygį, kad gyventojai galėtų apsirūpinti visomis būtinomis gėrybėmis. Šis rodiklis paprastai transformuojamas į perkamąją galią, kuri tiksliau matuoja gyvenimo standartą, nes įvertina atskirų prekių kainas įvairiose šalyse, pavyzdžiui, kiek mokama už kepalą duonos. Tai p_4 rodiklis.

Vienas iš bazinių rodiklių yra *gyventojų raštingumas*. Raštingumo turinys įvairiais visuomenės ekonominės raidos etapais nebuvo vienas. Besivystančiose šalyse raštingais laikomi mokantys skaityti, išsivysčiusiose – mokantys skaityti ir rašyti. UNESCO generalinė konferencija 1958 m. rekomendavo gyventojų surašymo metu raštingais laikyti mokančius skaityti, suprasti skaitomą tekstą ir galinčius parašyti rašinėlių iš kasdienio gyvenimo.

Baziniai rodikliai p_1 , p_2 , p_3 standartizuojami, t y. transformuojami į intervalą $[0;1]$ pagal formulę:

$$i_p = \frac{p_{fakt} - p_{min}}{p_{max} - p_{min}} \quad (6.1)$$

Čia: p_{fakt} – faktinė rodiklio reikšmė tam tikroje šalyje;
 p_{min} ir p_{max} – atitinkamai, minimali ir maksimali kiekvieno rodiklio reikšmė (ribinė reikšmė).

Standartizuojant p_4 indeksą (realaus BVP vienam gyventojui JAV doleriais) pagal perkamosios galios paritetą skaičiuojama šio rodiklio natūrinio logaritmo reikšmė:

$$i_{BVP} = \frac{\ln p_{4fakt} - \ln p_{4min}}{\ln p_{4max} - \ln p_{4min}}. \quad (6.2)$$

Standartizuojant rodiklius naudojamos šios ribinės reikšmės (žr. 6.1 lentelę):

6.1 lentelė. Ribinės rodiklių reikšmės

Rodiklis	p_{min}	p_{max}
Tikėtina gyvenimo trukmė (metais)	25	85
Suaugusių gyventojų raštingumo lygis (procentais)	0	100
Besimokančio jaunimo dalis (procentais)	0	100
Realus BVP <i>per capita</i> pagal perkamosios galios paritetą (JAV dol.)	100	40000

Skaičiuojant rodiklių p_1, p_2, p_3, p_4 standartines reikšmes pagal (6.1) išraišką, taikomos šios formulės:

- standartinis tikėtinos gyvenimo trukmės gimstant rodiklis:

$$i_{p1} = \frac{p_{1fakt} - 25}{85 - 25}; \quad (6.3)$$

- standartinis suaugusių gyventojų raštingumo rodiklis:

$$i_{p2} = \frac{p_{2fakt} - 0}{100 - 0}; \quad (6.4)$$

- standartinis dalies besimokančio jaunimo rodiklis:

$$i_{p3} = \frac{p_{3fakt} - 0}{100 - 0}; \quad (6.5)$$

- pasiekto išsilavinimo lygio indeksas apibendrina du standartinius rodiklius:

$$i_{apib} = \frac{(2 \cdot i_{p2} + i_{p3})}{3}; \quad (6.6)$$

- standartinis realaus BVP *per capita* pagal perkamosios galios paritetą rodiklis:

$$i_{p4} = \frac{\ln p_{4,fakt} - \ln 100}{\ln 40000 - \ln 100} = \frac{\ln p_{4,fakt} - 4.605}{10.597 - 4.605} = \frac{\ln p_{4,fakt} - 4.605}{5.992}; \quad (6.7)$$

- socialinio išsivystymo indeksas apibendrina tris standartinius bazinius rodiklius:

$$HDI = \frac{(i_{p1} + i_{apib} + i_{p4})}{3}. \quad (6.8)$$

Lyčių lygybės raidos indeksas GDI vertina kiekvienos šalies vyrų ir moterų gyvenimo trukmę, raštingumą ir pajamas., Tarptautiniuose palyginimuose naudojami abu indeksai, o valstybių rikiavimas pagal *HDI* ir *GDI* parodo lyčių išsivystymo netolygumą.

6.3. Gyventojų skurdo indeksas

Gyventojų skurdo indeksas apibūdina gyventojų vargingumo lygį. Juo vertinamos trys svarbiausios žmogaus gyvenimo sritys: gyvenimo trukmė, žinios ir aprūpinimas ekonominiais ištekliais. Šis indeksas skaičiuojamas pagal metodikas, sukurtas atskirai besivystančioms šalims ir išsivysčiusioms šalims.

Besivystančių šalių gyventojų skurdo indeksas (HPI₁) skaičiuojamas įvertinant šiuos bazinius rodiklius:

- p_1 – procentinė dalis gyventojų, kurių tikėtina gyvenimo trukmė neviršys 40 metų;
- p_2 – neraštingų suaugusių gyventojų procentinė dalis;
- p_3 – gyventojų aprūpinimas ekonominiais ištekliais; jį apibūdina du rodikliai:
 - p_{31} – procentinė dalis gyventojų, kurie neturi galimybės apsirūpinti tinkamos kokybės geriamu vandeniu,
 - p_{32} – procentinė dalis vaikų iki 5 metų amžiaus, kurių svoris daug mažesnis už normalųjį.

Bendras gyventojų aprūpinimo ekonominiais ištekliais rodiklis yra šių dviejų mažesnių bazinių rodiklių vidurkis:

$$p_3 = \frac{p_{31} + p_{32}}{2}. \quad (6.9)$$

Besivystančių šalių gyventojų skurdo indeksas HPI_1 skaičiuojamas pagal formulę:

$$HPI_1 = \sqrt[3]{\frac{1}{3}(p_1^3 + p_2^3 + p_3^3)}. \quad (6.10)$$

Išsivysčiusių šalių gyventojų skurdo indeksu (HPI_2) apibendrinama keturių bazinių rodiklių teikiama informacija:

- p_1 – procentinė dalis gyventojų, kurių tikėtina gyvenimo trukmė neviršys 60 metų;
- p_2 – procentinė dalis neraštingų suaugusių gyventojų;
- p_3 – procentinė dalis gyventojų, gyvenančių žemiau skurdo ribos, (tai gyventojai, kurių pajamos mažesnės nei 50 proc. visų gyventojų pajamų medianos);
- p_4 – procentinė dalis bedarbių, neturinčių darbo ilgą laikotarpį (12 mėn. ir daugiau).

Išsivysčiusių šalių gyventojų skurdo indekso reikšmė apskaičiuojama pagal formulę:

$$HPI_2 = \sqrt[3]{\frac{1}{4}(p_1^3 + p_2^3 + p_3^3 + p_4^3)}. \quad (6.11)$$

6.4. Rodiklių konstravimo pagrindumas

Pagrįstumas yra vienas iš esminių integruotų rodiklių reikalavimų. Pagrįstumas nagrinėjamas trimis aspektais: turinio pagrįstumas, konstravimo schemos pagrįstumas ir prognozavimo pagrįstumas. Konstravimo pagrįstumas turėtų būti patikrintas statistiniais metodais. Šie metodai taikomi tiriant bazinių rodiklių ryšius ir atrenkant didžiausios įtakos turinčius rodiklius. Baziniai rodikliai gali būti siejami tiek tiesiniu, tiek ir netiesiniu būdu. Siejimo būdui ir bazinių rodiklių svoriui nustatyti naudojami koreliacinės, regresinės ir faktorinės analizės metodai. Sudarytų

integruotų kriterijų prognozavimo objektyvumas tikrinamas gautus rezultatus lyginant su žinomomis reikšmėmis. Kadangi apskaičiavus integruotus rodiklius dažniausiai pateikiamos ne gautos reikšmės, o jų rangai, rodiklių konstravimo pagrįstumą galima patikrinti ranginės koreliacinės analizės metodais.

Tiriant ryšį, kuris gali egzistuoti tarp dviejų didėjimo ar mažėjimo tvarka išrikiuotų kintamųjų, galima įvesti ryšio krypties, taip pat ir koreliacijos sąvoką. Naudojant ranginius ryšio matavimus, kintamieji išrikiuojami, jiems suteikiami rangai, ir ***ranginiai koreliacijos koeficientai skaičiuojami tarp šių rangų, o ne tarp pradinių kintamųjų reikšmių***. Ranginiai koreliacijos koeficientai kinta nuo -1 iki $+1$.

Tarkime, kad žinome keliolikos valstybių dviejų rodiklių reikšmes, pavyzdžiui, socialinės pažangos indeksą *HDI* ir tikėtiną gyvenimo trukmę gimstant p_1 . Kuo valstybės socialinio išsivystymo lygis yra aukštesnis, tuo tikėtina gyvenimo trukmė gimstant turėtų būti ilgesnė. Valstybes galima išrikiuoti pagal rodiklį p_1 . Vieta, kurią užims valstybė *N* šiame sąraše, ir bus šios valstybės rangas pagal tikėtiną gyvenimo trukmę gimstant. Po to valstybes galima išrikiuoti pagal socialinės pažangos indeksą *HDI* ir užrašyti naujus valstybių rangus. Jeigu *HDI* sukonstruotas teisingai, tai valstybių rikiavimas pagal *HDI* bus panašus į rikiavimą pagal tikėtiną gyvenimo trukmę gimstant.

6.5. Spirmeno (Spearman) koreliacijos koeficientas

Ryšio stiprumui tarp dviejų kintamųjų, kurių skirstinys nėra normalusis, apskaičiuoti 1904 metais Spirmenas pasiūlė ranginį kriterijų, kuris ryšio stiprumui įvertinti naudoja ne pačias kintamųjų reikšmes, o jų rangus. Spirmeno ranginis koreliacijos koeficientas yra naudojamas tiriant ryšį, kuris gali egzistuoti tarp dviejų ranginių kintamųjų.

Principas, kuriuo remiantis skaičiuojamas Spirmeno koreliacijos koeficientas, yra labai paprastas. Skaičiuojant šį koreliacijos koeficientą palyginami dviejų sekų rangai. Šiuo tikslu surandami skirtumai tarp rangų, jie keliami kvadratu ir sudedami. Tada į skaičiavimą įtraukiama papildomų koeficientų tam, kad koeficiento reikšmės kistų nuo -1 (tai rodo labai stiprų neigiamą ryšį) iki $+1$, t. y. iki labai stipraus teigiamo ryšio. Jei koeficientas lygus nuliui, tai rodo, kad statistinio ryšio nėra.

Jeigu skirtumas tarp vienos rangų poros yra D_i , tai visų porų skirtumų kvadratų suma $\sum_{i=1}^n D_i^2$, o Spirmeno koreliacijos koeficientas r_s apskaičiuojamas pagal formulę:

$$r_s = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n^3 - n}. \quad (6.12)$$

Čia n – lyginamų elementų skaičius.

Pastaba. Šis paprastas Spirmeno koeficiento skaičiavimo būdas naudojamas tada, kai visi rangai skirtingi, t.y., kai tarp stebimų požymių nebuvo pasikartojančių reikšmių.

Pavyzdys. Du ekspertai įvertino 10 jiems pateiktų projektų ir sunumeravo juos pagal efektyvumą. Pirmasis ekspertas suteikė rangus $R_x=(1; 2; 3; 4; 5; 6)$. Antrasis ekspertas suteikė rangus $R_y=(3; 1; 2; 5; 6; 4)$. Reikia nustatyti, ar ekspertai panašiai įvertino projektus, t.y., ar tarp jų sudarytų rangų sekų yra ryšys.

Iš pirmo žvilgsnio sunku pasakyti, kiek šie įverčiai susiję vienas su kitu. Norint nustatyti, ar egzistuoja statistinis ryšys tarp šių reikšmių stulpelių, apskaičiuojamas Spirmeno koeficientas:

$$r_s = 1 - \frac{6[(-2)^2 + 1^2 + 1^2 + (-1)^2 + (-1)^2 + 2^2]}{6^3 - 6} \approx 0,66.$$

Tai vidutinio stiprumo koreliacija. Gauta reikšmė parodo, kad nors rangai nėra karto nesutapo, specialistų nuomonės yra panašios. Tai leidžia teigti, kad projektai įvertinti objektyviai.

Sutampančių rangų skaičiavimas. Rikiuojant kintamuosius eilės tvarka, galimi atvejai, kai du ar daugiau elementų turi tą pačią reikšmę. Tada kiekvienam tokiam elementui suteikiamas rangas, kuris yra vieno elementų rangų aritmetinis vidurkis. Pavyzdžiui, jei trys vienodi elementai užima atitinkamai penktą, šeštą ir septintą vietas, tai jiems suteikiamas vienodas rangas yra $\frac{5+6+7}{3} = 6$, o jei antrą ir trečią vietas,

tai jų rangas yra $\frac{2+3}{2} = 2,5$. Taigi rangai gali būti ne tik sveikieji, bet ir trupmeniniai skaičiai. Jeigu yra sutampančių rangų, Spirmeno koreliacijos koeficientas skaičiuojamas pagal apibendrintą formulę:

$$r_s = \frac{\sum_{i=1}^n \left(R_{xi} - \frac{n+1}{2} \right) \left(R_{yi} - \frac{n+1}{2} \right)}{\sqrt{\sum_{i=1}^n \left(R_{xi} - \frac{n+1}{2} \right)^2} \sqrt{\sum_{i=1}^n \left(R_{yi} - \frac{n+1}{2} \right)^2}}. \quad (6.13)$$

Čia $R_{x1}, R_{x2}, \dots, R_{xn}$ – pirmosios sekos rangai, o $R_{y1}, R_{y2}, \dots, R_{yn}$ – antrosios sekos rangai.

Pavyzdys. Dešimt įmonių pirmiausia buvo įvertintos pagal valdymo efektyvumą, po to – pagal gautą pelną. Šio įvertinimo rezultatas – dvi rangų sekos: $R_x = (1; 2,5; 2,5; 4,5; 4,5; 6,5; 6,5; 8; 9,5; 9,5)$ ir $R_y = (1; 2; 4,5; 4,5; 4,5; 4,5; 8; 8; 8; 10)$.

Apskaičiavę pagal formulę (6.13) gauname rezultatą $r_s = 0,917$. Tai labai stipri koreliacija, vadinasi, abu požymiai labai glaudžiai susiję tarpusavyje.

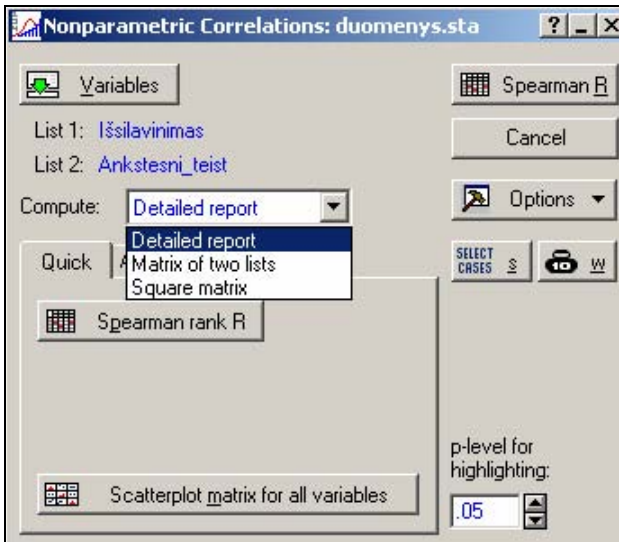
Jeigu skaičiuotume pagal formulę (6.12) tai būtų reikšmė $r_s = 0,892$.

Skaičiuojant Spirmeno koreliacijos koeficientą STATISTICA programa, meniu skyriuje *Statistics* pasirenkamas punktas *Nonparametrics*. Šio punkto analizės sąrašo trečioji eilutė yra *Correlations (Spearman, Kendall tau, gamma)*. Duomenų lange (6.1 pav.) nurodome kintamuosius ir rezultatų pateikimo būdą. Kintamųjų reikšmes gali sudaryti arba pradiniai duomenys, arba jų rangai. Jei įrašomi pradiniai duomenys, programa rangus apskaičiuoja be papildomų nurodymų.

Skaičiavimo rezultatai pateikiami lentelėje (6.2 lentelė). Šioje lentelėje pateikta:

- stebimų sekų ilgis (*Valid N*);
- apskaičiuota Spirmeno koreliacijos koeficiento reikšmė (*Spearman R*),
- reikšmė *t* kriterijaus, kuriuo tikrinama hipotezė H_0 , kad koreliacija statistiškai nereikšminga,

- stebimasis reikšmingumo lygmuo p -level.



6.1 pav. STATISTICA programos Spirmeno koreliacijos koeficiento pradinų duomenų langas. Laukelyje „Compute“ pasirenkama pageidautinas rezultatų pateikimo būdas.

6.2 lentelė. Spirmeno koreliacijos koeficiento skaičiavimo STATISTICA programa rezultatai

<i>Spearman Rank Order Correlations</i>				
	<i>Valid</i>	<i>Spearman</i>		
	N	R	t(N-2)	p-level
Išsilavinimas&Ankstesni_teist.	57	-0,211504	-1,60487	0,114251

Jeigu išmatuotų reikšmių yra $n \geq 10$, tai Spirmeno koreliacijos koeficiento skirstinys yra apytiksliai Stjudento t skirstinys su $n-2$ laisvės laipsniais. Hipotezę dėl šio koreliacijos koeficiento reikšmingumo galima tikrinti t kriterijumi:

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}}. \quad (6.14)$$

Čia: r_s – apskaičiuotas Spirmeno koeficientas; $(n-2)$ – Stjudento skirstinio laisvės laipsniai.

Gavus skaičiavimo rezultatus tenka priimti sprendimą, t. y. nutarti, ar hipotezę H_0 (H_0 : koreliacijos koeficientas lygus nuliui) atmesti. Sprendimo priėmimo taisyklė pasirenkama pagal skaičiavimo rezultatų pateikimo būdą (ar skaičiavimams naudota statistinė programa, ar kita skaičiavimo priemonė).

Sprendimo priėmimas, skaičiavimui naudojant statistinę programą

Kai skaičiavimui naudojama statistinė programa, ar atmesti hipotezę H_0 sprendžiama pagal stebimąjį reikšmingumo lygmenį p -level.

Taisyklė: Jei stebimojo reikšmingumo lygmens p -level reikšmė mažesnė už pasirinktą reikšmingumo lygmenį α (p -level $< \alpha$), tai hipotezė H_0 atmetama ir daroma išvada, kad koreliacija yra reikšminga. Jei p -level reikšmė didesnė ar tokia pati, kaip reikšmingumo lygmuo α (p -level $\geq \alpha$), tai koreliacija nereikšminga, o kintamieji nepriklausomi.

Sprendimo priėmimas, kai statistinio kriterijaus reikšmė apskaičiuota ir stebimasis reikšmingumo lygmuo p -level nežinomas

Šiuo atveju sprendimo priėmimo taisyklė tokia. Parenkamas reikšmingumo lygmuo ($\alpha=0,05$), statistinėje lentelėje randama Stjudento skirstinio su $(n-2)$ laisvės laipsniais $\alpha/2$ lygmens kritinė reikšmė $t_{\alpha/2}(n-2)$, čia: n – imties dydis. Hipotezė H_0 atmetama, jei apskaičiuota t kriterijaus absoliuti reikšmė didesnė nei $t_{\alpha/2}(n-2)$. Statistikos $t_{\alpha/2}(n-2)$ reikšmes galima apskaičiuoti statistinių duomenų analizės programų tikimybių pasiskirstymo skaičiuotuvais.

6.6. Rodiklių skaičiavimo pavyzdžiai

6.1 pavyzdys

Apskaičiuokime Lietuvos socialinės pažangos indeksą HDI , remdamiesi 1996 metų duomenimis (žr. 6.3 lentelę).

6.3 lentelė. Lietuva, 1996 metai

Rodiklis	<i>p_{fakt}</i>
Tikėtina gyvenimo trukmė (metais)	$p_{vvrū}=64,47; p_{moterų}=76,0; p_1=70,3$
Suaugusių gyventojų raštingumo lygis (proc.)	$p_2=99,8$
Besimokančio jaunimo dalis (proc.)	$p_3=48$ (EUROSTAT duomenys)
Realus BVP <i>per capita</i> (JAV dol.) palyginamosiomis kainomis	$p_4=1701$

Komponentinių indeksų reikšmės:

$$i_{p1} = \frac{70,3 - 25}{85 - 25} = 0,755;$$

$$i_{p2} = \frac{99,8 - 0}{100 - 0} = 0,998;$$

$$i_{p3} = \frac{48 - 0}{100 - 0} = 0,48;$$

$$i_{apib} = [2 \cdot 0,998 + 0,48] : 3 = 0,825;$$

$$i_{p4} = \frac{\ln(1701) - 4,605}{5,992} = 0,473.$$

Galutinė indekso reikšmė:

$$HDI = \frac{0,755 + 0,825 + 0,473}{3} = 0,683.$$

Palyginimui apskaičiuokime Danijos socialinės pažangos indeksą pagal 1995 metų duomenis (6.4 lentelė).

6.4 lentelė. Danija, 1995 metai

Rodiklis	<i>p_{fakt}</i>
Tikėtina gyvenimo trukmė (metais)	$p_1=75,3$
Suaugusių gyventojų raštingumo lygis (proc.)	$p_2=99$
Besimokančio jaunimo dalis (proc.)	$p_3=89$
Realus BVP <i>per capita</i> (JAV dol.) palyginamosiomis kainomis	$p_4=21963$

Komponentinių indeksų reikšmės:

$$i_{p1} = \frac{75,3 - 25}{85 - 25} = 0,838;$$

$$i_{p2} = \frac{99 - 0}{100 - 0} = 0,99;$$

$$i_{p3} = \frac{89 - 0}{100 - 0} = 0,89;$$

$$i_{apib} = (2 \cdot 0,99 + 0,89) : 3 = 0,957;$$

$$i_{p4} = \frac{\ln(21983) - 4,605}{5,992} = 0,900.$$

Galutinė indekso reikšmė

$$HDI = \frac{0,838 + 0,957 + 0,900}{3} = 0,898.$$

Lietuvos ir Danijos socialinės pažangos indeksų santykis yra $0,898/0,683=1,315$. Vadinasi, kad Danijoje žmogiškasis potencialas išvystytas 1,315 karto daugiau nei Lietuvoje.

6.2 pavyzdys

Apskaičiuokime Kongo gyventojų skurdo indeksą (HPI_I), remdamiesi šiais duomenimis (6.5 lentelė):

6.5 lentelė. Kongas, 1990 metai

Rodiklis	P_{fakt}
Gyventojai, kurių tikėtina gyvenimo trukmė neviršys 40 metų (proc.)	32,0
Neraštingų suaugusių gyventojų (proc.)	25,1
Gyventojai, neturintys galimybės apsirūpinti tinkamos kokybės geriamu vandeniu (proc.)	66,0
Vaikai iki 5 metų amžiaus, kurių svoris mažesnis už normą (proc.)	24,0

Apskaičiuokime $p_3 = (66 + 24) : 2 = 45$:

$$\text{tada } HPI_1 = (1/3(32^3 + 25,1^3 + 45^3))^{1/3} = 35,98$$

Vadinasi, 35,98 proc. Kongo gyventojų gyvena skurde ir neturi galimybių gyventi patenkinamomis sąlygomis.

6.3 pavyzdys

Apskaičiuokime Australijos gyventojų skurdo indeksą (HPI_2), remdamiesi šiais duomenimis (žr. 6.6 lentelę):

6.6 lentelė. Australija, 1990 metai

Rodiklis	<i>P_{fakt}</i>
Gyventojai, kurių tikėtina gyvenimo trukmė neviršys 60 metų (proc.)	9,0
Neraštingų suaugusių gyventojų (proc.)	17,0
Gyventojų, kurių pajamos mažesnės nei 50 proc. visų gyventojų pajamų mediana (proc.)	12,9
Bedarbiai, nedirbantys ilgiau nei 1 metus (proc.)	2,6

Australijos gyventojų skurdo indeksas

$$HPI_2 = [1/4(9^3 + 17^3 + 12,9^3 + 2,6^3)]^{1/3} = 12,5$$

Vadinasi, 12,6 proc. Australijos gyventojų gyvena žemiau skurdo ribos.

Užduotys

6.1. 6.7 lentelėje pateikti duomenys apibūdina devintojo dešimtmečio vidurio gyvenimo kokybę kai kuriose išsivysčiusiose pasaulio šalyse.

Remdamiesi šiais duomenimis:

1. Apskaičiuokite kiekvienos valstybės socialinės pažangos indeksą *HDI*.
2. Apskaičiuokite Spirmeno koreliacijos koeficiento reikšmes, apibūdinančias indekso *HDI* priklausomumą nuo bazinių rodiklių. Kuris bazinis rodiklis turi didžiausios įtakos indeksui *HDI* išsivysčiusių šalių grupėje?

6.7 lentelė. Išsivysčiusių pasaulio šalių gyvenimo kokybė devintojo dešimtmečio viduryje

Valstybė	Tikėtina gyvenimo trukmė gimstant (metais)	Suaugusių gyventojų raštingumo lygis (proc.)	Besimokančio jaunimo dalis (proc.)	BVP <i>per capita</i> (JAV dol.) palyginamosiomis kainomis
Belgija	76,9	99,0	86,0	21548
Austrija	76,7	99,0	87,0	21322
Ispanija	77,7	97,1	90,0	14789
Graikija	77,9	96,7	82,0	11636
Airija	76,4	99,0	88,0	17590
Italija	78,0	98,1	73,0	20174
Olandija	77,5	99,0	91,0	19876
Naujoji Zelandija	76,6	99,0	94,0	17267
Norvegija	77,6	99,0	92,0	22427
Kanada	79,1	99,0	100,0	21916
Islandija	79,2	99,0	83,0	21064
JAV	76,4	99,0	96,0	26977
Prancūzija	78,7	99,0	89,0	21176
Japonija	79,9	99,0	78,0	21930

6.2. 6.8 lentelėje pateikti duomenys apibūdina devintojo dešimtmečio vidurio gyvenimo kokybę besivystančių šalių grupėje.

6.8 lentelė. Besivystančių pasaulio šalių gyvenimo kokybė

Valstybė	Tikėtina gyvenimo trukmė gimstant (metais)	Suaugusių gyventojų raštingumo lygis (proc.)	Besimokančio jaunimo dalis (proc.)	BVP <i>per capita</i> (JAV dol.) palyginamosiomis kainomis
1	2	3	4	5
Argentina	72,6	96,2	79,0	8498
Brazilija	66,6	83,3	61,0	5928
Venesuela	72,3	91,1	67,0	8090
Singapūras	77,1	91,1	68,0	22604

1	2	3	4	5
Kolumbija	70,3	91,3	69,0	6347
Tailandas	69,5	93,8	55,0	7742
Malaizija	71,4	83,5	61,0	9572
Meksika	72,1	89,6	67,0	6769
Turkija	68,5	82,3	60,0	5516
Omanas	70,7	59,0	60,0	9383
Kuveitas	75,4	78,6	58,0	23848
Honkongas	79,0	92,2	67,0	22950
Čilė	75,1	95,2	73,0	9930
Bahreinas	72,2	85,2	84,0	16751
Fidži	72,1	91,6	78,0	6159

Remdamiesi šiais duomenimis:

1. Apskaičiuokite kiekvienos valstybės socialinės pažangos indeksą *HDI*.
2. Apskaičiuokite Spirmeno koreliacijos koeficiento reikšmes, apibūdinančias socialinio pažangos indekso *HDI* priklausomumą nuo bazinių indeksų. Kuris bazinis indeksas turi didžiausios įtakos indeksui *HDI* analizuojamų šalių grupėje?

6.3. 6.9 lentelėje pateikti duomenys apibūdina gyventojų skurdo lygį besivystančių pasaulio šalių grupėje.

1. Apskaičiuokite besivystančių šalių gyventojų skurdo indeksą. Kurioje šalyje didžiausia skurdo indekso reikšmė, kurioje – mažiausia?
2. Tarp gautos skurdo indekso reikšmės ir jo bazinių rodiklių apskaičiuokite porinius koreliacijos koeficientus ir patikrinkite jų reikšmingumą. Palyginkite koeficientus. Kuris bazinis rodiklis turi didžiausios įtakos skurdo lygiui, o kuri – mažiausios?
3. Parenkite rezultatų analizę.

6.9 lentelė. Besivystančiose pasaulio šalių gyventojų skurdo lygis

Valstybė	Gyventojų, kurių gyvenimo trukmė neviršys 40 metų (proc.)	Raštingumo lygis (proc.)	Gyventojų, neaprupintų kokybišku vandeniu (proc.)	Vaikai per mažo svorio iki 5 m. (proc.)
Kamerūnas	26,0	63,4	50,0	14,0
Gana	23,0	64,5	35,0	27,0
Nigerija	31,0	57,5	50,0	36,0
Mauritanija	29,0	37,7	26,0	23,0
Madagaskaras	21,0	45,8	66,0	34,0
Gvinėja-Bisau	42,0	54,9	41,0	23,0
Gvinėja	38,0	35,5	54,0	26,0
Malis	36,0	31,0	34,0	27,0
Etiopija	34,0	35,5	75,0	48,0
Burkina-Faso	38,0	19,2	22,0	30,0
Siera-Leone	36,0	13,6	52,0	36,0
Mozambikas	38,0	40,1	37,0	27,0

6.4. 6.10 lentelėje pateikti duomenys apibūdina gyventojų skurdo lygį išsivysčiusių pasaulio šalių grupėje 1995 m.

6.10 lentelė. Gyventojų skurdo lygis išsivysčiusiose pasaulio šalyse 1995 m.

Valstybė	Gyventojai, kurių gyvenimo trukmė neviršys 60 metų (proc.)	Raštingumo lygis (proc.)	Gyventojai žemiau skurdo ribos (proc., pagal ES standartą)	Ilgalaikė bedarbystė (proc. nuo visos darbo jėgos)
1	2	3	4	5
Airija	9,0	99,0	11,1	7,6
Kanada	9,0	99,0	11,7	1,3
Prancūzija	11,0	99,0	7,5	4,9
Norvegija	9,0	99,0	6,6	1,3
JAV	13,0	99,0	19,1	0,5
Suomija	11,0	99,0	6,2	6,1
Olandija	9,0	99,0	6,7	3,2
Japonija	8,0	99,0	11,8	0,6

1	2	3	4	5
Švedija	8,0	99,0	6,7	1,5
Ispanija	10,0	97,5	10,4	13,0
Vokietija	12,0	99,0	7,5	2,0
Jungtinė Karalystė	9,0	99,0	13,5	3,8
Danija	11,0	99,0	5,9	4,0

1. Apskaičiuokite išsivysčiusių šalių gyventojų skurdo indeksą. Kurios šalies skurdo indekso reikšmė didžiausia, kurios – mažiausia?
2. Apskaičiuokite gautos skurdo indekso reikšmės ir jo bazinių rodiklių porinius koreliacijos koeficientus ir patikrinkite jų reikšmingumą. Palyginkite koeficientus tarpusavyje. Kuris bazinis rodiklis turi didžiausios įtakos skurdo lygiui, o kuris – mažiausios?
3. Parašykite rezultatų analizę.

6.5. 6.11 lentelėje pateikti duomenys apie Lietuvos apskritis. Nustatykite, ar šios požymių mažėjimo tvarka surašytos apskritys susijusios tarpusavyje. Apskaičiuokite Spirmeno koreliacijos koeficiento reikšmę. Patikrinkite hipotezę, kad šis koeficientas lygus nuliui.

6.11 lentelė. Lietuvos apskritys, surašytos rodiklių mažėjimo tvarka (2000 m.)

Užregistruota nusikaltimų 10 000 gyv.	Gyventojų skaičius 1 kv. km	Nedarbo lygis	Automobilių sk. 10 000 gyv.
Klaipėda	Kaunas	Šiauliai	Marijampolė
Vilnius	Vilnius	Tauragė	Klaipėda
Panevėžys	Klaipėda	Marijampolė	Tauragė
Kaunas	Šiauliai	Panevėžys	Alytus
Telšiai	Marijampolė	Telšiai	Telšiai
Šiauliai	Telšiai	Alytus	Utena
Marijampolė	Panevėžys	Utena	Panevėžys
Tauragė	Alytus	Klaipėda	Šiauliai
Utena	Tauragė	Kaunas	Kaunas
Alytus	Utena	Vilnius	Vilnius

7. TEISINĖ STATISTIKA

„Mokslas panašus į detektyvinį apsakymą. Visi faktai tvirtina tą pačią hipotezę, tačiau galiausiai pasirodo, kad teisinga yra visai kita.“

F. Frankas

Ilgą laiką teisinė statistika buvo vadinama moraline arba dorovine. Šį pavadinimą 1833 m. pirmą kartą pavartojo prancūzų statistikas A. Guerry. Moralinės statistikos apibrėžimų buvo daug, o tipiškas to meto apibrėžimas skambėtų taip: „Moralinė statistika tiria masinius moralinius žmonių gyvenimo reiškinius“. Tačiau tiriamus reiškinius kiekvienas tyrėjas suprato savaip. Tai galėjo būti nusikaltimai, baudmės, santuokos, skyrybos, gimimai (tarp jų ir nesantuokiniai), abortai, palikti naujagimiai, venerinės ligos, alkoholizmas, religinės apeigos, sektantizmas, prostitucija, savižudybės ir pan. Ypatingą vietą moralinės statistikos struktūroje užėmė kriminalinė statistika. Tai buvo susiję ne su ypatingu tyrėjų dėmesiu kriminaliniams duomenims, o su tuo, kad nusikaltimų, savižudybių, nužudymų ir pan. apskaita buvo ženkliai geresnė. XX a. pradžioje terminą „moralinė statistika“ buvo pradėta kritikuoti dėl neapibrėžtumo ir buvo siūloma jį keisti į keletą terminų, būtent: „kriminalinė statistika“, „savižudybių statistika“, „alkoholio vartojimo statistika“ ir pan. Šiuo metu vartojamas terminas „teisinė statistika“.

7.1. Teisiniai statistiniai rodikliai

Atskirose valstybėse, taip pat ir visame pasaulyje tobulinamos duomenų rinkimo metodikos ir renkami šie duomenys:

- bendras užregistruotų nusikaltimų skaičius;
- asmenys, įvykdę nusikaltimus;
- asmenys, patraukti baudžiamojon atsakomybėn;
- nuteistieji;

- išteisintieji;
- nuteistieji laivės atėmimu;
- nusikaltimų aukos;
- nusikaltimų pasekmės;
- nusikaltimų padaryta žala;
- nusikaltimų priežastys ir sąlygos;
- nusikaltimų motyvai;
- nusikaltimų padarymo būdai;
- administraciniai teisės pažeidimai;
- administracinės baudos;
- administracinės teisės pažeidėjai;
- teismų ir teisėsaugos organų veikla;
- penitencinė statistika.

Bendras nusikalstamumo lygio rodiklis yra *užregistruotų nusikaltimų skaičius*.

Didžiausią grėsmę socialiniam saugumui sudaro kriminaliniai nusikaltimai. Kriminologijos teorijoje į nusikaltimo priežastis, konkretų nusikaltimą ir nusikaltėlio asmenybę paprastai žiūrima kaip į daugiau ar mažiau savarankiškas, nors ir susijusias, problemas. Nusikalstamo elgesio motyvai dažnai suvedami į asmenybės problemą. Praktikoje teisėsaugos institucijoms sudėtinga tiksliai atskirti nusikaltimo priežastis, nusikaltėlio asmenybę, nusikaltimo motyvus ir įvykdytą nusikaltimą. Šios srities statistinės analizės tikslas – išmatuoti ir išaiškinti ryšius, kurie egzistuoja tarp požymių, užrašytų kriminalinėje apskaitoje, ir tarp socialinės bei ekonominės aplinkos, kurioje gyveno ir veikė nusikaltimą įvykdęs asmuo. Šiuolaikinės teisinės statistikos stiprioji pusė yra ta, kad ji neapsiriboja vien kriminaline apskaita, naudojami ir kiti socialinės statistikos rodikliai. Ypatingą vietą čia užima nusikalstamumo lygio ir pagrindinių socialinių bei ekonominių rodiklių tarpusavio ryšiai. Teisinė statistika nagrinėja teisinius reiškinius ir procesus tam, kad būtų galima atpažinti juose vykstančių procesų šablonus, tendencijas, raidos, apibrėžtos konkrečių vietos ir laiko sąlygų, dėsningumus.

Paskutiniajame šimtetyje daug kartų buvo bandoma palyginti atskirų valstybių kriminalinę statistiką. Šiuo tikslu buvo atliekami tyrimai, kurie aprėpė atskiras labiau išsivysčiusias šalis. Buvo stengiamasi sukurti rodiklius, leidžiančius nors šiokiu tokiu atžvilgiu palyginti atskirų valstybių su skirtingais baudžiamaisiais ir procesiniais įstatymais, tei-

smine praktika ir kitomis statistiškai svarbiomis sąlygomis, nusikalstamumo lygį. Po Antrojo pasaulinio karo šia veikla užsiėmė Jungtinės Tautos. Jau 1949 m. išleistame statistikos metraštyje buvo pateikta statistinių žinių iš daugiau kaip 20 šalių apie nusikalstamumo dinamiką karo ir pokario metais.

1950 m. JT priėmė rezoliuciją apie būtinybę kas 5 metai kviesti JT tarptautinius kongresus siekiant užkirsti kelią nusikalstamumui, kadangi daugelyje pasaulio regionų nuolat augantis nusikaltimų skaičius pradėjo kelti grėsmę valstybių raidai ir normaliam žmonių gyvenimui. 1976 m. JT pasaulio valstybėms išsiuntinėjo anketas, kuriose turėjo būti įrašytas bendras asmenų, 1970–1975 m. įvykdžiusių nusikaltimus, skaičius, taip pat jų pasiskirstymas pagal amžių ir lytį. Be to, valstybės turėjo pranešti bendrą užregistruotų nusikaltimų skaičių ir pasiskirstymą pagal dešimt nusikaltimų rūšių:

- tyčiniai nužudymai,
- pavojingi pasikėsinimai į asmens gyvybę ir turta;
- lytiniai nusikaltimai;
- žmonių grobimai;
- plėšimai;
- vagystės;
- sukčiavimai;
- neteisėta prekyba narkotinėmis medžiagomis;
- narkotikų vartojimas;
- alkoholinių gėrimų vartojimas.

Atsakymus atsiuntė 64 valstybės, ir jie buvo pateikti pirmojoje JT statistinėje ataskaitoje. Rengiantis antrajai statistinei ataskaitai, sudarytame klausimyne buvo išskirtos trys duomenų grupės:

1. Statistiniai duomenys apie užregistruotus nusikaltimus.
2. Statistiniai duomenys ir kokybinė informacija apie teisėtvarkos sistemos funkcionavimą.
3. Žinios apie nusikalstamumo prevencijos strategiją.

Ši kartą atsakymus pateikė 80 valstybių, ir jie buvo išspaudinti antrojoje JT statistinėje ataskaitoje. Devintajame JT kongrese 1995 m. jau 100 valstybių pateikė duomenis, kurie buvo išspausdinti ketvirtojoje JT statistinėje ataskaitoje.

Statistiniai duomenys, kuriuos pateikia JT, turi daugelį trūkumų, iš kurių didžiausias – tai šių duomenų patikimumas. JT duomenų netikrina

ir spausdina tai, ką pateikia valstybės. Kai kurios valstybės pateikia neišsamius duomenis. Be to, duomenys penkerius metus vėluoja.

Surinkta teisinė statistinė informacija leidžia JT sekti bendras nusikalstamumo tendencijas pasaulyje bei atskiruose regionuose, prognozuoti nusikalstamumą, kurti rekomendacijas baudžiamosios teisės unifikavimui, kurti ir siūlyti kovos su nusikalstamumu strategiją ir prevencines priemones.

Šiuo metu visame pasaulyje pastebimos dvi bendros teisinės statistikos raidos tendencijos:

1. Teisiškai reikšmingų faktų statistinės apskaitos tobulinimas, detalizavimas ir tikslinimas.
2. Kriminalinės apskaitos dominavimas.

Lietuvoje teisinės statistikos raidai trukdė buvęs duomenų įslaptinimas, prastai ir negausiai surenkami duomenys. Tyrimams trukdo ne tiek duomenų stygius, kiek juos jungiančių grandžių nebuvimas. Pavyzdžiui, pateikiamose ataskaitose apie nuteistųjų skaičių galima rasti informacijos atskirai apie nuteistų asmenų amžių ir atskirai apie bausmės terminą, tačiau nepateikiama visa informacija, iš kurios būtų galima spręsti, koks bausmės termino ir nuteistų asmenų amžiaus santykis. Didžiausias teisinės statistikos informacijos srautas gaunamas iš valstybinių teisėsaugos institucijų. Dalis ataskaitų patenka į Lietuvos statistikos departamentą. Apskričių statistiką renka vietos savivaldybės. Į statistikos įstaigas patenka statistinė informacija apie užregistruotus nusikaltimus ir jų rūšis, apie asmenis, įvykdžiusius baudžiamuosius ar administracinius pažeidimus, apie teisėsaugos sistemos veikimą, apie numatomas prevencines priemones.

7.2. Rodiklių tipai

Absoliutūs rodikliai (laiko eilutės). Pagrindiniai teisinės statistikos rodikliai yra absoliutūs apibendrinti duomenys, apie kurių dydį ir kitimą galima spręsti palyginus juos su ankstesnių metų duomenimis (7.1 lentelė).

Šie absoliutūs rodikliai (laiko eilutės) yra analogiški demografinės statistikos duomenims, ir jie analizuojami tais pačiais metodais, kurie pateikti demografinėje statistikoje.

7.1 lentelė. Lietuvos statistikos departamento pateikti duomenys apie bendrą užregistruotų nusikaltimų skaičių

Metai	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Užregistruota nusikaltimų	37056	44984	56615	60378	58634	60819	68053	75816	78149	77108	82370

Sugrupuoti duomenys. Dažnai teisiniai statistinių ataskaitų duomenys yra sugrupuoti. Pavyzdžiui, duomenys apie nuteistųjų pasiskirstymą pagal amžių, bausmės terminą, nusikaltimą, išsilavinimą dažnai pateikiami taip, kaip yra 7.2 lentelėje.

7.2 lentelė. Pravieniškių BR PDK 2000 m. pateikti duomenys apie nuteistųjų pasiskirstymą pagal amžių

Nuteistųjų amžius	Nuteistųjų skaičius
Iki 21 metų (imtina)	78
Nuo 21 metų iki 30 metų (imtina)	196
Nuo 30 metų iki 40 metų (imtina)	49
Nuo 40 metų iki 50 metų (imtina)	15
Nuo 50 metų iki 60 metų (imtina)	6
Vyresni kaip 60 metų	1

Sugrupuoti duomenys yra priskiriami prie klasikinių *rangų* skalės kintamųjų, gaunamų siejant duomenis į klases. Taip galima sieti ne tik amžiaus, bet ir kitus kiekybinius kintamuosius. Pavyzdžiui, pajamas galima skirstyti intervalais iki 500 Lt, nuo 501 iki 1000 Lt, nuo 1000 iki 2000 Lt, per 2000 Lt ir pan. Nagrinėjant šiuos duomenis svarbu sąrašė nurodyta eilė. Šiuos duomenis galima palyginti tarpusavyje.

Jeigu reikia nustatyti ryšį (koreliaciją) su kitu ranginio tipo kintamuoju, naudojami *ranginiai koreliacijos koeficientai*, kurie yra pateikti integruotų socialinės raidos rodiklių skyriuje. Jei ranginio kintamojo skalę sudaro 15 ir daugiau pozicijų, tokį kintamąjį galima nagrinėti kaip kiekybinį.

Ryšiai tarp ranginių kintamųjų gali būti tiriami naudojant nominalinius kriterijus, tačiau šie kriterijai neatspindi informacijos, kurią suteikia rangai. Todėl ranginių kintamųjų ryšių analizei geriau naudoti Spirmeno ar Kendalo ranginius koreliacijos koeficientus, o ne χ^2 kriterijų.

Pagrindinės sugrupuotų duomenų charakteristikos

Paprasčiausias būdas rasti sugrupuotų duomenų *medianą* būtų visas į intervalą patekusias reikšmes pakeisti vidurinėmis intervalo reikšmėmis ir pritaikyti medianos skaičiavimo formulę.

Sugrupuotų duomenų aritmetinis vidurkis:

$$\bar{x} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{i=1}^m f_i}, \quad (7.1)$$

čia: f_i – grupės elementų skaičius, m – grupių skaičius.

Sugrupuotų duomenų dispersija:

$$s^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 f_i}{\sum_{i=1}^m f_i}. \quad (7.2)$$

7.3. Koreliacinė analizė. Pirsono koreliacijos koeficientas

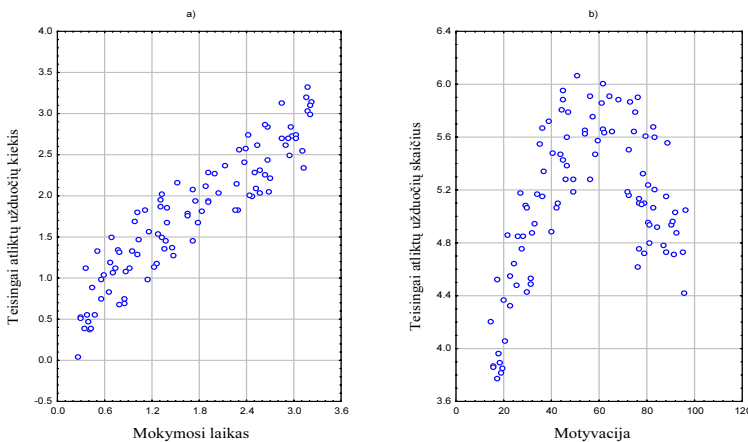
Prognozuojant teisinius ar demografinius rodiklius, būtina įvertinti šių rodiklių ryšius su kitais socialiniais bei ekonominiais rodikliais, arba rodiklių kitimo priklausomumą nuo laiko. Požymiai, apibūdinantys tiriamą objektą, dažnai būna susiję tarpusavyje, priklauso vienas nuo kito, sąlygoja vienas kitą ir kinta laike. Ypatingą vietą čia užima priežastinumas, kai vienas ar keli tarpusavyje susiję reiškiniai, vadinami priežastimi, sukelia kitus reiškinius, vadinamus pasekme arba rezultatu. Priežastiniai ryšiai gali būti nusakomi tik tikimybiškai, kadangi pakeitę tam tikrą sąlygą kaskart gausime vis kitą rezultatą. Tokia priežastinio ryšio forma, kai priežastis nusako pasekmę ne vienareikšmiškai, o tik su tam tikra tikimybe, yra neišsami ir vadinama priežastine priklausomybe. Terminas „koreliacija“ pradinė reikšmė – tarpusavio ryšys (Oxford Advanced Learner’s Dictionary of Current English, 2002). Kalbant apie

koreliaciją, vartojami du terminai „koreliacinis ryšys“ ir „koreliacinė priklausomybė“. Šie terminai dažnai vartojami kaip sinonimai. Tuo tarpu koreliacinis ryšys tarp požymių kartais rodo ne šių požymių tarpusavio ryšį, o jų priklausomumą nuo kažkokio trečio požymio ar kelių požymių, kurie tyrime nenagrinėjami. Priklausomybė reiškia įtaką, o ryšys – bendrą ar darnų dviejų požymių kitimą. Koreliaciniai ryšiai neparodo, kuris požymis yra priežastis, o kuris pasekmė. Jie liudija, kad vieno požymio kitimas sukelia kito požymio kitimą, tačiau nerodo, ar kitimo priežastis yra viename iš požymių, ar yra už tyrimo ribų. Kartais galima stebėti ar kontroliuoti kurio nors požymio kitimą ir sekti, kaip jo kitimas veikia kitus požymius. Toks požymis vadinamas *nepriklausomu*. Požymiai, kurie, mūsų nuomone, kinta veikiant nepriklausomiems požymiams, vadinami *priklausomais*. Priklausomų ir nepriklausomų kintamųjų bendras kitimas vadinamas *koreliacine priklausomybe*.

Dažniausiai tyrime sunku nustatyti, kuris požymis yra priklausomas, o kuris nepriklausomas. Kadangi žodžiu „priklausomybė“ suprantama įtaka, geriau vartoti neutralų terminą „ryšys“.

Koreliaciniai ryšiai skiriasi pagal formą, kryptį ir stiprumą.

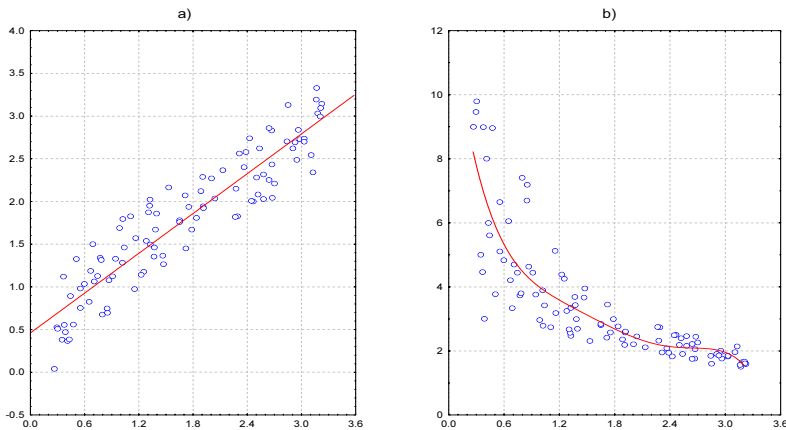
Pagal *formą* koreliaciniai ryšiai skirstomi į *tiesinius* ir *kreivinius* (*netiesinius*). Tiesiniai gali būti, pavyzdžiui, tarp mokymuisi sugaišto laiko ir patikrinimo metu teisingai atliktų užduočių skaičiaus (7.1 pav., a).



7.1 pav. Tiesinis (a) ir netiesinis (b) koreliacinis ryšys

Netiesiniai ryšiai gali būti stebimi, pavyzdžiui, tarp motyvacijos lygio ir teisingai atliktų užduočių skaičiaus. Didėjant motyvacijai, teisingai atliktų užduočių daugėja, pasiekia optimalų motyvacijos lygį, kurį atitinka maksimalus efektyvumas; toliau motyvacijai didėjant efektyvumas pradeda mažėti (7.1 pav., b)

Pagal *kryptį* koreliaciniai ryšiai yra *teigiami* („tiesioginiai“) ir *neigiami* („atvirkštiniai“). Esant teigiamam tiesiniam ryšiui, vieno požymio reikšmėms didėjant, didėja ir kito požymio reikšmės (7.2 pav., a). Esant neigiamam ryšiui, vieno požymio reikšmėms didėjant, kito požymio reikšmės mažėja (7.3 pav., b)



7.2 pav. Teigiamas (tiesinis) ir neigimas (netiesinis) koreliacinis ryšys

Kai žinome, kad egzistuoja statistinis ryšys (koreliacija) tarp dviejų kintamųjų, tai galime teigti, kad jų reikšmių kitimas yra susijęs. Tačiau remdamiesi koreliacija negalime pasakyti, kuris kintamasis sukelia šiuos pokyčius. Koreliacija nerodo kitimo priežasties. Žinodami, kad kintamieji X ir Y yra koreliuoti, dažniausiai negalime pasakyti, ar Y priklauso nuo X , ar X priklauso nuo Y , ar jie abu yra susiję su dar kažkoku faktoriumi, kuris ir sukelia jų abiejų kitimą. Kai X ir Y yra statistiškai nepriklausomi, tai žinodami X kitimą negalime prognozuoti Y kitimo, tiksliau, žinios apie X kitimą nepagerina Y numatymo. Kuo ryšys stipresnis,

tuo tiksliau pagal X kitimą galima prognozuoti Y kitimą. Kiekybinių kintamųjų ryšio stiprumą galima išmatuoti Pirsono (*Pearson*) koreliacijos koeficientu. Didelės šio koeficiento reikšmės, nežiūrint ar jos teigiamos, ar neigiamos, atitinka tai, ką vadiname stipria koreliacija, o mažos reikšmės – silpna koreliacija. Jei koreliacija yra nereikšminga, tai nereiškia, kad koreliacijos koeficientas tiksliai lygus nuliui, tačiau jo reikšmė yra arti nulio. Sociologai ir psichologai dažnai tiria daug kintamųjų, ir jiems svarbu nustatyti, *kuris* iš šių kintamųjų yra labiausiai susijęs su turimu kintamuoju. Kiekybinių kintamųjų tyrimui naudojami keli skirtingi koreliacijos koeficientai. Dažniausiai naudojamas Pirsono koreliacijos koeficientas r . Šis koeficientas yra apskaičiuojamas kaip dviejų sekų reikšmių porų sandauga, kai iš kiekvienos reikšmės yra atimamas vidurkis ir skirtumas padalijamas iš standartinių nuokrypių sandaugos:

$$r = \frac{\frac{1}{n-1} \sum (x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}; \quad (7.3)$$

čia \bar{x} ir \bar{y} atitinkamai yra stebėjimų x ir y vidurkiai, o s_x ir s_y – standartiniai nuokrypiai.

Koreliacijos koeficientas turi šias savybes:

- koreliacijos koeficientas yra dydis, kintantis nuo -1 iki 1 ;
- kai $r = 1$, tai visi taškai (x_i, y_i) yra tiesėje, kurios krypties koeficientas yra teigiamas;
- kai $r = -1$, tai visi taškai (x_i, y_i) yra tiesėje, kurios krypties koeficientas yra neigiamas;
- $r = 0$, kai kintamieji yra tiesiškai nepriklausomi.

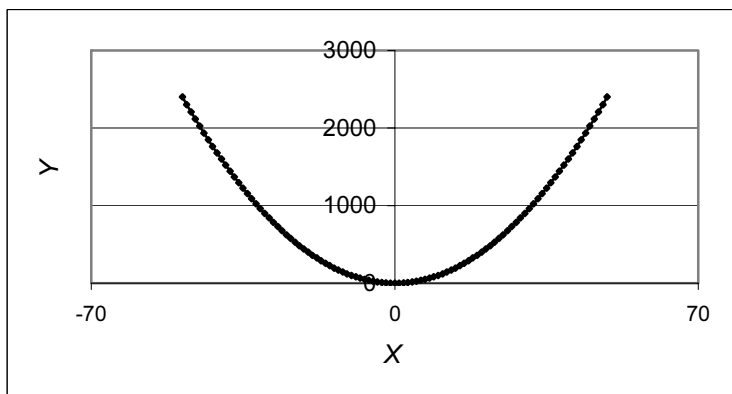
Ilgoms imtims (t. y. dideliems n) yra nusistovėjęsios tam tikros taisyklės, kokią koreliaciją laikyti stipria (apie tai – 7.3 lentelėje).

Šios taisyklės tinka pradiniam koreliacijos stiprumui įvertinti ir tik esant didelėms imtims (kai n daugiau už 30). Ženkliai tiksliau įvertinsime koreliacinį ryšį patikrinę hipotezę dėl koreliacijos koeficiento lygybės nuliui. Todėl kitame tyrimo etape turime įsitikinti, kad koreliacijos koeficientas iš esmės skiriasi nuo nulio.

7.3 lentelė. Koreliacijos stiprumo interpretacija

Koreliacijos koeficiento reikšmė	Interpretacija
Nuo 0,9 iki 1,0 arba nuo -0,9 iki -1,0	Labai stipri koreliacija
Nuo 0,7 iki 0,9 arba nuo -0,7 iki -0,9	Stipri koreliacija
Nuo 0,5 iki 0,7 arba nuo -0,5 iki -0,7	Vidutinė koreliacija
Nuo 0,3 iki 0,5 arba nuo -0,3 iki -0,5	Silpna koreliacija
Nuo 0,3 iki -0,3	Koreliacija nereikšminga

Pastaba. Būtina pabrėžti, kad r nurodo tik *tiesinį* ryšį. Jei apskaičiavote koreliacijos koeficientą $r = 0$, nereikia skubėti daryti išvados, kad jokio ryšio tarp kintamųjų apskritai nėra. Gali pasitaikyti ir netiesiniai ryšys, o iš to, kad r artimas nuliui, galima daryti išvadą, kad nėra tinkamos tiesės pagal šiuos duomenis.



7.3 pav. Netiesinis ryšys, kurio koreliacijos koeficientas $r=0$

Daugelyje socialinių tyrimų ryšys tarp požymių visai pagrįstai gali būti vaizduojamas tiese. Tačiau nereikia, kad negali pasitaikyti išimčių.

Kartais gali tecti apskaičiuoti stiprią koreliaciją tarp kintamųjų, kurių nesieja jokie loginiai ryšiai. Tai vadinamoji *nesąmoninga* koreliacija. Visiškai atsitiktinai kažkas gali apskaičiuoti ryšį tarp studentų skaičiaus ir skaičiaus žuvų, sugautų vandenynuose, bei gauti reikšmę, pavyzdžiui, $r = 0.98$. Tačiau kartais tokie statistiniai ryšiai atskleidžia trečią faktorių, nuo kurio priklauso abu kintamieji.

Koreliacijos koeficiento reikšmingumo tikrinimas. Suformuluokime statistinę hipotezę.

Nulinė (pagrindinė) hipotezė H_0 teigia, kad koreliacijos koeficientas lygus nuliui.

Alternatyvioji hipotezė H_1 teigia, kad koreliacijos koeficientas nelygus nuliui.

Hipotezės H_0 teisingumui patikrinti naudosime t kriterijų:

$$t = r \sqrt{\frac{n-2}{1-r^2}}. \quad (7.4)$$

Čia: r – apskaičiuota koreliacijos koeficiento reikšmė, n – išmatuotų reikšmių kiekis.

Parinkime reikšmingumo lygmenį. Tegul bus reikšmingumo lygmuo $\alpha=0,05$. Hipotezė H_0 atmetama, jei absoliuti t reikšmė viršija dydį $t_{\alpha/2(n-2)}$, kuris yra Stjudento skirstinio su $(n-2)$ laisvės laipsniais $\alpha/2$ lygmens kritinė reikšmė.

Dalinis koreliacijos koeficientas. Kai nagrinėjami kintamieji X ir Y kartu su kitais kintamaisiais (Z, W, \dots), tai būna dažnas atvejis, kai kintamieji X ir Y kinta dėl to, kad kinta kiti kintamieji. Dalinis koreliacijos koeficientas parodo tiesinio ryšio tarp X ir Y stiprumą tuo atveju, kai pašalinama kitų kintamųjų įtaka. Trijų kintamųjų atveju dalinį koreliacijos koeficientą ρ galima apskaičiuoti pagal formulę:

$$\rho_{XY(Z)} = \frac{r_{XY} - r_{xz}r_{YZ}}{\sqrt{(1-r_{XZ}^2)(1-r_{YZ}^2)}}; \quad (7.5)$$

čia r_{XY} – koreliacijos koeficientas tarp kintamųjų X ir Y .

Determinacijos koeficientas. Vertinant modelio tinkamumą analizuojamos liekanos ir skaičiuojamas *determinacijos* koeficientas. Statistikai determinacijos koeficientą naudoja dažniau nei koreliacijos koeficientą. Determinacijos koeficientas parodo, kuri vieno požymio bendro kitimo dalis gali būti paaiškinta kito požymio reikšmių kitimu. Determinacijos koeficientas R^2 tiesinės priklausomybės atveju yra lygus koreliacijos tarp kintamųjų (požymių) X ir Y koeficiento kvadratui. Apskritai šis koeficientas interpretuojamas taip:

$$R^2 = 1 - \frac{s_{\varepsilon}^2}{s_y^2} = \frac{\text{Variacijos dalis, kurią paaiškina modelis}}{\text{Visa variacija}}. \quad (7.6)$$

Čia: s_{ε}^2 – liekanų (ε) dispersija, s_y^2 – stebimų kintamojo (Y) reikšmių dispersija.

Paprastosios tiesinės regresijos atveju determinacijos koeficientas yra koreliacijos koeficiento kvadratas:

$$R^2 = r_{xy}^2. \quad (7.7)$$

Kai koreliacijos koeficientas $r = 0$, tai $R^2 = 0$. Taip ir turi būti, kadangi šiuo atveju X kitimas neveikia Y kitimo ir jo paaiškinti negali. Analogiškai, kai $r = \pm 1,0$, tai $R^2 = 1,0$. Vadinasi, Y kitimą visiškai galima paaiškinti X reikšmių kitimu. Nagrinėjant kitus atvejus, pavyzdžiui, kai $r = \pm 0,3$, tai $R^2 = 0,09$; tik 9 procentus vieno kintamojo kitimo galima paaiškinti kito požymio kitimu. Kai $r = \pm 0,5$, tai $R^2 = 0,25$ (25 procentai Y kitimo paaiškinami X kitimu), o kai $r = \pm 0,8$, tai $R^2 = 0,64$ (paaiškinami 64 procentai Y kitimo).

7.4. Kriminalinių ir socialinių bei ekonominių rodiklių ryšiai ir jų analizės metodai

Šiuolaikinė teisinė statistika neapsiriboja vien kriminalinių reiškinių apskaita. Informacinė visuomenės raida sąlygoja ne tik geresnę statistinę nusikalstamumo apskaitą, bet ir sudaro geras galimybes įvairiai informacijai apie atskirų valstybių socialinę bei ekonominę ir kriminalinę padėtį kaupti. Norint geriau suprasti ir valdyti kriminalinę situaciją bei sukurti efektyvią prevencinės veiklos strategiją, būtina įvertinti ryšius, kurie egzistuoja tarp požymių, užrašytų kriminalinėje apskaitoje, ir tarp socialinės bei ekonominės aplinkos, kurioje gyvena ir veikė nusikaltimą įvykdęs asmuo. Statistika leidžia ne tik įvertinti nusikalstamumo kitimo tempus, bet ir nustatyti konkrečiai valstybei ar regionui būdingas nusikalstamumo priežastis. Čia ypatingą vietą užima bendrųjų šalies ekonominių rodiklių (bendrojo vidaus produkto, kaupimo ir vartojimo, gyventojų užimtumo, alkoholinių gėrimų gamybos ir vartojimo ir t.t.) įtakos nusikalstamumo tendencijoms analizė.

Galima analizuoti nusikalstamumo ryšius su pačiais įvairiausiai šalies socialinės bei ekonominės raidos rodikliais. Tačiau daugelis iš jų glaudžiai susiję tarpusavyje, todėl analizei reikia atrinkti pačius įtakingiausius ir reikšmingiausius.

Ekonominę šalies situaciją geriausiai apibūdina dalis bendrojo vidaus produkto (BVP) palyginamosiomis kainomis vienam gyventojui.

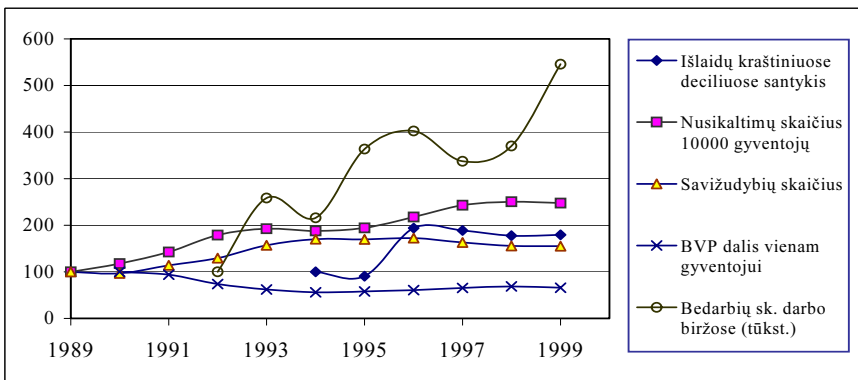
Socialinių visuomenės teisingumo lygį atspindi ekonominis visuomenės narių susisluoksniavimas. Daugeliu tyrimu XXVIII–XX a. buvo įrodyta, kad šis socialinis bei ekonominis faktorius yra ypač kriminogeninis. Jau Platonas, norėdamas užkirsti kelią nusikalstamumui, siūlė tiesiog nustatyti skurdo ir turto santykį 1:4. Todėl nusikalstamumo analizei labai svarbus rodiklis – turtingiausio decilio (t.y. 10 proc. turtingiausių šeimų narių) vartojimo išlaidų (vienam asmeniui) ir skurdžiausio decilio (t.y. 10 proc. neturtingiausių šeimų narių) vartojimo išlaidų santykis. Labai svarbus rodiklis – gyventojų užimtumas. Didelis bedarbių skaičius, ypač jaunimo, visuomenei visada kėlė didžiausią rūpestį. Todėl nedarbo lygis ar bedarbių skaičius yra vienas iš nagrinėtinų rodiklių.

7.4 lentelė. Lietuvos ekonominiai bei socialiniai ir nusikalstamumo rodikliai

METAI	Nusikaltimų skaičius 10000 gyv.	Savižudybės	BVP vienam gyventojui palyginamosiomis kainomis (tūkstančiais Lt)	Bedarbiai, užregistruoti darbo biržoje (tūkstančiais)	Bedarbių skaičius gyventojų apklausų duomenimis (tūkstančiais)	Išlaidų kraštiniuose deciliuose santykis
1989	84,16	1000				
1990	99,2	969	11,166			
1991	120,0	1142	10,478			
1992	150,5	1294	8,250	30,96		
1993	162,0	1572	6,933	79,95		
1994	158,0	1703	6,272	66,81	347,2	4,49
1995	163,6	1694	6,488	112,5	347,1	4,08
1996	183,3	1723	6,804	124,5	317,4	8,72
1997	204,5	1632	7,307	104,5	257,2	8,49
1998	210,9	1554	7,687	114,5	232,4	7,98
1999	208,3	1552	7,374	168,9	281,8	8,07

Išnagrinėkime, kaip Lietuvoje kito nusikaltimų, savižudybių, bendrojo vidaus produkto (BVP) vienam gyventojui palyginamosiomis kainomis, bedarbių, užregistruotų darbo biržoje bei apklausiant gyventojus, skaičius ir mėnesinių išlaidų (kraštiniuose deciliuose) santykis vienam šeimos nariui (7.4 lentelė).

Šiame pavyzdyje nagrinėsime, kokios priežastys formuoja tiriamo rodiklio (šiuo atveju – nusikaltimų skaičiaus) lygį ir kokia konkrečios priežasties įtaka. Būtent rodiklio kitimo priklausomybės nuo aplinkos sąlygų tyrimas sudaro koreliacijos teorijos turinį. Bendros tendencijos tampa aiškesnės, šiuos rodiklius pateikus grafiškai (7.4 pav.). Masteliams suvienodinti sudarysime dinamines eilutes, t.y. pirmųjų nagrinėjamų metų (dažniausiai 1989 m.) rodiklius prilyginsime 100, o visų kitų metų rodiklius apskaičiuosime pagal jų santykį su pirmųjų metų rodikliais.



7.4 pav. Lietuvos ekonominių bei socialinių ir nusikalstamumo rodiklių dinaminės eilutės

Ryšių tarp atskirų rodiklių stiprumui nustatyti apskaičiuosime porinius koreliacijos koeficientus, kurie parodo tiesinės priklausomybės tarp dviejų rodiklių stiprumą. Tačiau poriniai koreliacijos koeficientai gali parodyti ryšį ir tada, kai rodikliai nepriklauso vienas nuo kito, tačiau abu priklauso nuo kitų rodiklių. Tiesinio ryšio tarp dviejų rodiklių stiprumą, nepriklausomai nuo kitų rodiklių įtakos, parodo daliniai koreliacijos koeficientai. Todėl gilesniam ryšiui nustatyti reikia apskaičiuoti ir šiuos

koreliacijos koeficientus. Gauti skaičiavimų rezultatai pateikti 7.5 lentelėje.

7.5 lentelė. Nusikaltimų skaičiaus 10000 gyventojų priklausomybė nuo kitų rodiklių

Nusikaltimų skaičiaus 10000 gyventojų priklausomybė nuo:	Porinis koreliacijos koeficientas	Dalinis koreliacijos koeficientas
BVP vienam gyventojui palyginamosiomis kainomis (tūkst. Lt)	0,984	0,975
Bedarbių, užregistruotų darbo biržose skaičius (tūkstančiais)	0,64	0,45
Išlaidų kraštiniuose deciliuose santykio	0,84	0,715

Koreliacinės analizės išvados. 7.4 lentelėje bedarbių skaičių apibūdina dvi skirtingos duomenų sekos. Pirmoji sudaryta remiantis darbo biržos duomenimis, antroji – statistikos departamento duomenimis, gautais apklausus gyventojus. Šie duomenys nepanašūs, priklausomumas tarp jų labai silpnas (porinis koreliacijos koeficientas yra -0,37). Duomenų patikimumą galėtume įvertinti naudodamiesi įvairiais žinomais ryšiais, pvz.: lygindami juos su nusikaltimų skaičiaus kitimu, su bendrojo vidaus produkto (BVP) kitimo tendencijomis ir pan. Apskaičiavę porinius koreliacijos koeficientus gauname, kad bedarbių skaičiaus darbo biržose ir BVP vienam gyventojui nuo 1994 metų koreliacijos koeficiento įvertis yra 0,58, o BVP vienam gyventojui ir bedarbių skaičiaus (apklausų duomenimis) koreliacijos koeficientas yra -0,97. Tuo tarpu bedarbių skaičiaus darbo biržose ir nusikaltimų skaičiaus porinis koreliacijos koeficientas yra 0,64, o bedarbių skaičiaus (apklausų duomenimis) ir nusikaltimų skaičiaus -0,944. Taigi apklausų duomenimis nustatytas bedarbių skaičius nedidina nusikaltimų skaičiaus ir yra susijęs tik su šalies ekonomine situacija.

Nusikaltimų ir bedarbių skaičiaus darbo biržose dalinis koreliacijos koeficientas yra 0,45. Kadangi duomenų turime labai nedaug (bedarbių skaičius darbo biržose registruojamas nuo 1992 metų), reikia patikrinti, ar šis dalinis koreliacijos koeficientas yra reikšmingas. Taikydami 5 proc. reikšmingumo lygmenį gauname, jog ši ryšį keičia kitų kintamųjų

įtaka, todėl tirdami nusikaltimų skaičiaus kitimą, bedarbių skaičiaus galime nevertinti. Daliniai koreliacijos koeficientai tarp nusikaltimų skaičiaus ir BVP bei išlaidų kraštiniuose deciliuose santykio yra reikšmingi. Abu šie koeficientai teigiami, vadinasi, didėjant vienam rodikliui, didėja ir kitas. Iš to galima daryti išvadą, kad gerėjant Lietuvoje ekonominei situacijai nėra tendencijos nusikaltimų skaičiui mažėti.

7.5. Tiesinė regresinė analizė

Apskaičiavus koreliacijos koeficientą nustatomas ryšių egzistavimas, tačiau negalima pasakyti, kaip keisis, pavyzdžiui, dirbančių asmenų skaičius, keičiantis asmenų, kaltinamų padarius nusikaltimus, skaičiui, t. y. negalima to numatyti (prognozuoti). Prognozavimui dažnai taikomi regresijos modeliai. Regresinės analizės prognozės yra kiekybinės – regresijos funkcija aprašomas tiriamo rodiklio (kintamojo) vidurkio priklausomumas nuo kito rodiklio reikšmių kitimo. Svarbiausias regresinės analizės pranašumas yra tas, kad sudaro galimybę parinkti kintamuosius siejančią funkciją (sudaromas modelis).

Pirmiausia pasirenkamas regresijos modelio tipas – nusprendžiama, kokia priklausomybė sieja kintamuosius. Nagrinėsime *tiesinės* regresijos modelį, t. y. atvejį, kai vienas kintamasis nuo kito priklauso tiesiškai. Šis modelis dažnai taikomas socialiniuose moksluose, nes dažniausiai procesai visuomenėje kinta lėtai, ir taikant tiesinį modelį jie prognozuojami ganėtinai gerai.

Tiesinė kintamojo Y priklausomybė nuo kintamojo X reiškia lygtimi:

$$y = \alpha + \beta x + \varepsilon. \quad (7.8)$$

Šioje lygtyje α ir β yra nežinomi koeficientai (konstantos), o ε – atsitiktinė paklaida. Kaip atsiranda atsitiktinė paklaida? Ją gali sudaryti matavimo paklaidos arba neįvardytos ir netiriamos priežastys.

Parinę įvairias koeficientų α ir β reikšmes gausime tiesių šeimą. Kaip apskaičiuoti geriausiai tiesę atitinkančius koeficientus? Tai atliekama taikant matematinį metodą – *mažiausių kvadratų metodą*. Šiuo metodu gautos koeficientų α ir β reikšmės – įverčiai (jas žymėsime raidėmis a ir b), randamos minimizuojant reiškinį:

$$SSE = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2. \quad (7.9)$$

Sumą SSE minimizuojančios koeficientų α ir β reikšmės (įverčiai) nustatomos taip:

$$b = \frac{n \left(\sum_{i=1}^n x_i y_i \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right)^2}, \quad (7.10)$$

$$a = \frac{\sum_{i=1}^n y_i}{n} - b \frac{\sum_{i=1}^n x_i}{n}. \quad (7.11)$$

Lygtis $\hat{y}(x) = a + bx$, kurios koeficientai apibrėžti (7.10) ir (7.11) lygybėmis, yra regresijos tiesės lygtis, o $\varepsilon = y(x) - \hat{y}(x)$ yra liekamoji paklaida, arba liekana.

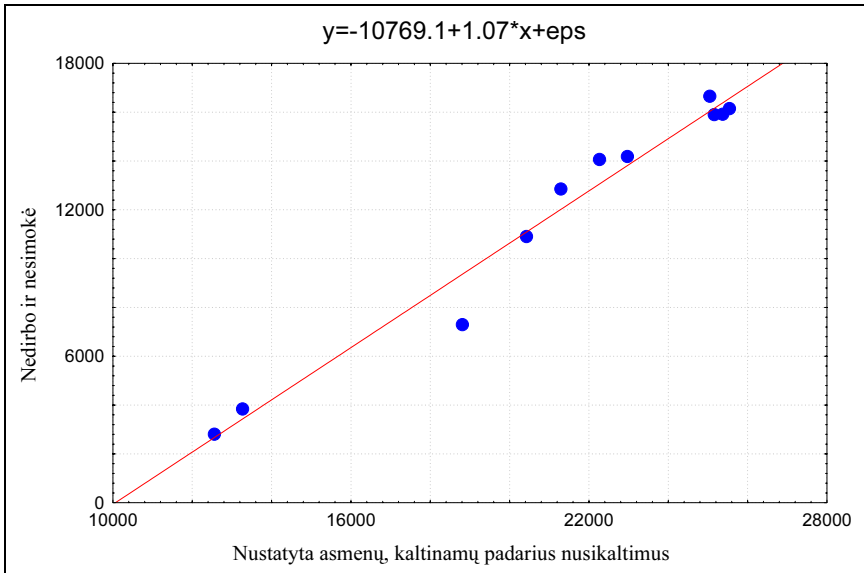
Apskaičiuokime tiesinės regresijos lygties koeficientus tuo atveju, kai x – asmenų, kaltinamų padarius nusikaltimus, skaičius, o y – dalis jų, kurie nedirbo ir nesimokė (7.6 lentelė).

7.6 lentelė. Pradiniai duomenys – asmenys, kaltinami padarę nusikaltimus, ir jų užimtumas

Metai	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Nustatyta asmenų, kaltinamų padarius nusikaltimus (x)	12556	13268	18810	20424	21290	22969	22269	25542	25373	25160	25046
Nedirbo ir nesimokė (y)	2808	3841	7300	10909	12854	14182	14066	16146	15920	15900	16653

Apskaičiavę gauname šias reikšmes: $a = -10769,1$, $b = 1,07$.

Parašome regresijos lygtį $\hat{y} = -10769,1 + 1,07x$ ir nubraižome diagramą (7.5 pav.).



7.5 pav. Regresijos tiesės diagrama

Apskaičiuoto koeficiento b reikšmė didesnė už vieną, taigi galima teigti, kad nedirbančių ir nesimokančių asmenų skaičius didėja greičiau nei bendras nusikaltimus padariusių asmenų skaičius. Remiantis šia lygtimi galima prognozuoti nedirbančių ir nesimokančių asmenų skaičių, pavyzdžiui, kai $x = 26000$, $\hat{y} = -10769,1 + 1,07 * 26000 = 17051$.

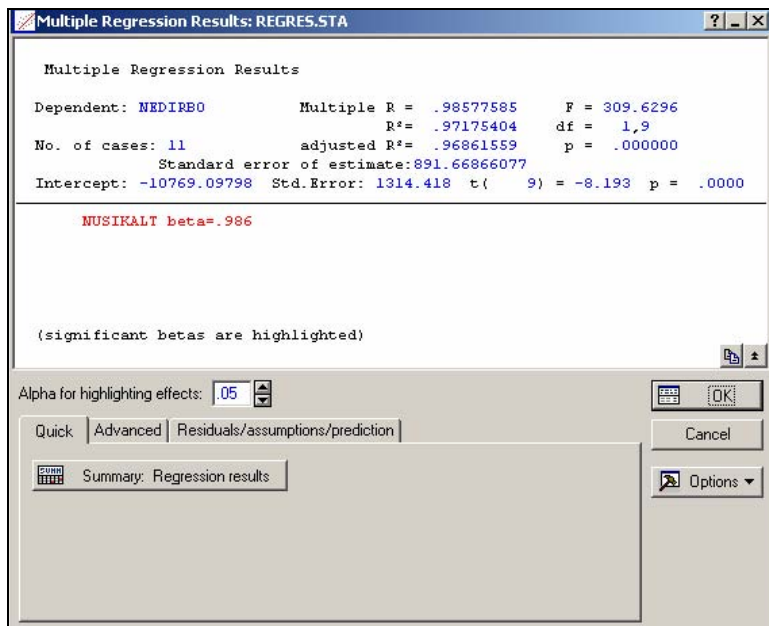
7.6. Tiesinė regresinė analizė *Statistica* programa

Regresijos koeficientų skaičiavimas, parinkto modelio tinkamumo turimiems duomenims tikrinimas užima daug laiko, todėl šiam darbui dažniausiai pasitelkiamos statistinės programos. Atliekant tiesinę regresinę analizę *Statistica* programa renkamos šios komandos:

Statistics

Multiple Regression.

Į duomenų langą įrašę priklausomą (*Dependent*) ir nepriklausomą (*Independent*) kintamuosius, gauname pradinį skaičiavimo rezultatą (7.6 pav.).



7.6 pav. Tiesinės regresijos pradinio skaičiavimo rezultatų dialoginis langas

Paspaudę klavišą *Summary:Regression results*, gauname bendrus tiesinės regresinės analizės skaičiavimo rezultatus (7.7 lentelė).

7.7 lentelė. Tiesinės regresinės analizės skaičiavimo rezultatai taikant STATISTICA programą

Regression Summary						
R=0,98577585 R²=0,97175404 Adjusted R²=0,96861559						
F(1,9)=309,63 p<0,00000 Std.Error of estimate: 891,67						
	BETA	St. Err. of BETA	B	St. Err. of B	t(9)	p-level
<i>Intercept</i>			-10769,098	1314,418	-8,193	1,829E-05
NUSIKALT	0,986	0,056	1,070	0,061	17,596	2,800E-08

7.7 lentelėje pateikti šie duomenys: standartizuotas regresijos koeficientas ($Beta$), regresijos lygties koeficientai (B), šių koeficientų standartiniai nuokrypiai, t kriterijus ir jo stebimasis reikšmingumo lygmuo.

Populiacijos parametrų įverčiai

Standartizuotas nedimensinis regresijos koeficientas **beta** nustatomas iš lygties:

$$beta = b_1 \frac{s_X}{s_Y}; \quad (7.12)$$

$Beta$ koeficientas yra pasvirimo koeficientas, apskaičiuotas tuo atveju, kai kintamieji X ir Y yra standartizuoti. Standartizavus duomenis, kintamųjų vidurkiai yra nulis, o standartinis nuokrypis – vienetas $\left(z_i = \frac{x_i - \bar{x}}{s}\right)$. Jeigu yra keli nepriklausomi kintamieji, jų standartizuoti regresijos koeficientai leidžia palyginti kiekvieno nepriklausomo kintamojo indėlį į priklausomo kintamojo prognozę.

Koeficientai a ir b , apskaičiuoti iš imties, skirsis nuo koeficientų α ir β , gaunamų iš populiacijos. Norint patikrinti hipotezes, ar iš imties duomenų apskaičiuotos šių koeficientų reikšmės nėra nulis, reikia žinoti jų skirstinius. Tiesinės regresijos atveju koeficientų α ir β skirstiniai yra normalieji su vidurkais a ir b . Koeficiento α standartinis nuokrypis (*standard error*) yra:

$$\sigma_\alpha = \sigma \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{(n-1)s_X^2}}. \quad (7.13)$$

Čia: s_X^2 yra atsitiktinio dydžio X dispersija.

Koeficiento β standartinis nuokrypis yra:

$$\sigma_\beta = \frac{\sigma}{\sqrt{(n-1)s_X^2}}. \quad (7.14)$$

Formulėse (7.13) ir (7.14) naudojama liekamosios paklaidos ε dispersija σ , kuri yra nežinoma, tačiau gali būti įvertinta. Jos įvertis apskaičiuojamas pagal formulę:

$$s^2 = \frac{\sum_{i=1}^n (y_i - a - bx_i)^2}{n-2}. \quad (7.15)$$

Kvadratinė šaknis iš s^2 vadinama *liekanų standartiniu nuokrypiu*, arba įvertinimo standartine paklaida (*standard error of the estimate*).

7.7 lentelės penktame stulpelyje *St. Err. of B* yra pateikti koeficientų α ir β standartiniai nuokrypiai, o lentelės viršuje – liekanų standartinis nuokrypis – *Std. Error of estimate*.

Hipotezių tikrinimas

Dažniausiai tikrinama hipotezė, kad tarp kintamųjų X ir Y nėra jokio tiesinio ryšio, t.y., kad koeficientas β yra nulis. Norint patikrinti šią hipotezę, naudojama statistika:

$$t = \frac{b}{s_\beta}. \quad (7.16)$$

Jei nulinė hipotezė, kad nėra jokio tiesinio ryšio, yra teisinga, ši statistika (7.14) turi Stjudento t skirstinį su $(n-2)$ laisvės laipsniais. Hipotezę, kad tiesė eina per koordinatų pradžią, t. y., kad koeficientas α lygus nuliui, galima patikrinti naudojantis kita statistika:

$$t = \frac{a}{s_\alpha}. \quad (7.17)$$

Statistika (7.17) taip pat turi Stjudento t skirstinį su $(n-2)$ laisvės laipsniais.

Šios abi t statistikos ir jų dvipusis stebimasis reikšmingumo lygmuo taip pat pateikti 7.7 lentelėje (du dešinieji stulpeliai – t ir p -level).

Parinkto modelio tikimas turimiems duomenims

Sudarius modelį, būtina iširti, ar šis modelis gerai tinka turimiems duomenims (*goodness of fit*). Šiame tyrime tikrinama, ar nepažeistos analizei būtinos sąlygos.

Tiesinio modelio tinkamumui tikrinti naudojami: koreliacijos koeficientas, determinacijos koeficientas ir pataisytasis determinacijos koeficientas.

Determinacijos koeficientu R^2 įvertinamas modelio tikimas visai populiacijai dažnai būna per daug optimistinis. Pataisytasis determinaci-

jos koeficientas R_a^2 tiksliau įvertina modelio tikimą duomenims:

$$R_a^2 = R^2 - \frac{p(1 - R^2)}{n - p - 1}; \quad (7.18)$$

čia: p yra nepriklausomų kintamųjų skaičius (nagrinėjamu atveju $p = 1$).

7.7 lentelės viršuje pateikta koreliacijos koeficiento reikšmė ($R=0,996$), determinacijos koeficiento reikšmė ($R^2=0,972$) ir pataisytojo (*adjusted*) determinacijos koeficiento reikšmė (*Adjusted* $R^2=0,969$). Iš gautų reikšmių galima spręsti, kad modelis puikiai tinka duomenims.

Didelė determinacijos koeficiento reikšmė reiškia, kad duomenys yra mažai nutolę nuo mažiausių kvadratų metodu gautos tiesės. Jei duomenys idealiai atitinka regresijos tiesės lygtį, t.y. visos y_i reikšmės patenka ant regresijos tiesės, determinacijos koeficientas yra 1. Jei determinacijos koeficientas artimas nuliui, tai regresijos tiesės lygtis visiškai netinka prognozei. Praktiškai taikant tiesinę regresinę analizę dažniausiai reikalaujama, kad $R^2 \geq 0,25$. Jei $R^2 < 0,25$, tiesinės regresijos modelio taikymas labai abejotinas.

7.7. Regresinė analizė ir kokybiniai kintamieji

Regresine analize tiriamą priklausomą kintamąjį dažnai veikia ne tik kiekybiniai kintamieji, kuriems matuoti naudojama tiksliai apibrėžta skalė (investicijos, pajamos, kainos, gyventojai, migracija, nedarbas, užterštumas, nusikaltimai), bet ir kintamieji, kurie iš esmės yra kokybiniai (lytis, rasė, religija, tautybė, socialiniai, ekonominiai ar politiniai pokyčiai). Pavyzdžiui: yra žinoma, kad dėstytojų moterų atlyginimai yra mažesni nei analogiškų pareigų dėstytojų vyrų, nors kiti atlyginimų dydžiai įtakos turintys faktoriai yra vienodi. Tai gali būti lyčių diskriminacijos rezultatas, tačiau, kad ir kokia būtų šio reiškinio priežastis, šiuo atveju kintamasis „lytis“ turi įtakos priklausomam kintamajam ir turi būti įtrauktas į nepriklausomų kintamųjų sąrašą.

Kokybiniai kintamieji dažniausiai parodo tam tikros savybės turėjimą ar neturėjimą. Tokia savybė gali būti vyras ar moteris, krikščionis ar nekrikščionis, juodaodis ar baltaodis ir t. t. Vienas iš būdų, padedančių apibrėžti tokius kintamuosius, yra dirbtinių kintamųjų, galinčių įgyti tik reikšmes 1 arba 0, sudarymas. Pavyzdžiui: 1 gali reikšti, kad asmuo

yra vyras, o 0 – moteris; 1 gali reikšti, kad asmuo turi aukštąjį išsilavinimą, o 0 – jo neturi ir pan. Kintamieji, kuriems priskiriamos reikšmės 0 arba 1, dar vadinami *fiktyviais* kintamaisiais arba *pseudokintamaisiais*. Taip pat jie vadinami ir kitaip – dichotominiais kintamaisiais, dvejetainiais kintamaisiais bei indikatoriniais kintamaisiais.

Nors dažniausiai fiktyviems kintamiesiems priskiriamos reikšmės 0 ir 1, tačiau tai nėra būtina ir privaloma sąlyga. Porą $(0,1)$ tiesine funkcija $(Y=a+bK$ ($b \neq 0$), kai a ir b yra konstantos, o K gali įgyti reikšmes 1 arba 0), galima transformuoti į kitokio pavidalo porą. Kai $K=1$, tai $Y=a+b$, o kai $K=0$, tai $Y=a$. Taigi pora $(0,1)$ tampa pora $(a;a+b)$. Pavyzdžiui, jei $a=1$ ir $b=2$, tai fiktyvus kintamasis gali būti $(1,3)$. Vadinasi, kokybiniai kintamieji neturi natūralios matavimo skalės.

Regresiniuose modeliuose fiktyvūs kintamieji naudojami taip pat, kaip ir kiekybiniai kintamieji. Iš tiesų regresiniame modulyje gali būti net keli pagalbiniai kintamieji, kurie iš esmės yra kokybiniai. Čia turėtų kilti klausimas: ar ne geriau kokybinio kintamojo išskirtus atvejus nagrinėti atskirai? Pavyzdžiui, jei nagrinėsime atlyginimo priklausomybę nuo darbo stažo ir lyties, galima būtų nagrinėti atskirai vyrų ir atskirai moterų atlyginimo priklausomybę nuo stažo. Šiuo atveju reikėtų vadovautis tokia taisykle: jeigu vyrų atlyginimo priklausomybę nuo darbo stažo apibūdina ta pati funkcija, kaip ir moterų atlyginimo priklausomybę nuo darbo stažo, tik atsiranda papildomas priedas, apibūdinamas kintamojo „lytis“ požymiu – „vyras“ (regresijos tiesės lygiagrečios), tai partartina naudoti fiktyvius kintamuosius. Kai kintamųjų priklausomybę apibūdina skirtingos funkcijos (su skirtingais koeficientais, regresijos tiesės nelygiagrečios), tai tokius atvejus reikia nagrinėti atskirai.

Kadangi šio tipo regresija dažnai naudojama socialiniuose tyrimuose, išnagrinėkime keletą jos taikymo pavyzdžių.

7.1 pavyzdys. Miestų ir rajonų gyventojų mirtingumo regresijos modelis

Pats paprasčiausias modelis – tik su vienu fiktyviu nepriklausomu kintamuoju:

$$y_i = \alpha + \beta k_i + \varepsilon_i; \quad (7.19)$$

čia: y_i – savivaldybių gyventojų mirtingumas, t. y. mirusiųjų skaičius 1000-iui gyventojų, o kodas k_i nustatomas taip:

$$k_i = \begin{cases} 0, & \text{jei savivaldybė yra miesto,} \\ 1, & \text{jei savivaldybė yra rajono.} \end{cases}$$

Šiuo atveju nagrinėjame tik dviejų kintamųjų regresijos modelį, kuriame vietoje kiekybinio kintamojo naudojamas kokybinis kintamasis k . Šis modelis ir dispersinės analizės ANOVA modelis yra analogiški.

Modelis (7.19) sudaro galimybę ištirti skirtumo tarp miestų ir rajonų gyventojų mirtingumo egzistavimą, tariant, kad visi kiti kintamieji, kaip gydytojų skaičius, vidutinis gyventojų amžius ir kt., yra pastovūs. Kadangi šiuo modeliu privalu tenkinti tas pačias sąlygas, kaip ir klasikiniais tiesinės regresijos modeliais, tai:

$$\text{vidutinis gyventojų mirtingumas miestuose } E(y_i | k_i = 0) = \alpha;$$

$$\text{vidutinis gyventojų mirtingumas rajonuose } E(y_i | k_i = 1) = \alpha + \beta,$$

vadinasi, koeficientas α yra mirusiųjų skaičius rajonuose 1000 gyventojų, $\alpha + \beta$ – mirusiųjų skaičius miestuose 1000 gyventojų, o jų skirtumas – koeficientas β .

Hipotezė, kad miestų ir rajonų gyventojų vidutinis mirtingumo skirtumas yra nulis ($H_0: \beta = 0$), tikrinama taip pat, kaip ir įprastinėje tiesinėje regresijoje, naudojant t kriterijų.

Šiame pavyzdyje išnagrinėkime penkiolikos Lietuvos miestų ir rajonų gyventojų mirtingumo, t. y. mirusiųjų skaičiaus 1000 gyventojų (žr. lentelę 7.8) 1999 m. duomenis.

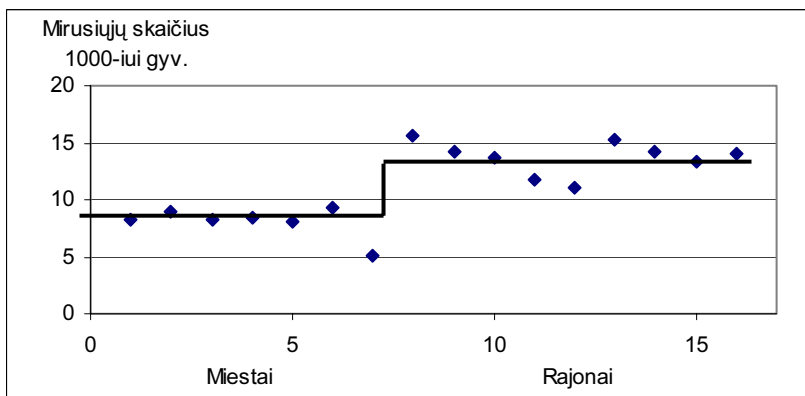
Pagal 7.8 lentelės duomenis apskaičiavus (7.19) lygties koeficientus, gauname: septynių miestų savivaldybių vidutinis gyventojų mirtingumas – 8,53 (a), o devynių rajonų savivaldybių gyventojų vidutinis mirtingumas – 13,67 ($a + b$). Regresijos lygtis:

$$\hat{y}_i = 8,35 + 5,14 \cdot k_i. \quad (7.20)$$

Abu (7.20) lygties koeficientai yra statistiškai reikšmingi, kai reikšmingumo lygmuo – 5 proc. Grafiškai lygtis pateikta 7.11 pav. Čia duomenys išrikiuoti ir sugrupuoti į dvi kategorijas – miestų ir rajonų savivaldybes. Regresijos funkcija yra laiptuota funkcija – skirtumas tarp miestų ir rajonų savivaldybių gyventojų mirtingumo yra 5,14 mirusiųjų 1000 gyventojų (7.7 pav.).

7.8 lentelė. Lietuvos miestų ir rajonų savivaldybių gyventojų mirtingumą ir sergamumą 1999 m. duomenys

Nr.	Savivaldybė	Kodas <i>k</i>	Mirusiųjų skaičius 1000 gyventojų	Susirgusių piktybiniais navikais skaičius 1000 gyventojų
1	Vilniaus m.	0	8,2	3,501
2	Kauno m.	0	9	4,136
3	Klaipėdos m.	0	8,2	3,926
4	Šiaulių m.	0	8,5	3,456
5	Panevėžio m.	0	8	3,496
6	Palangos m.	0	9,3	3,119
7	Anykščių r.	1	15,7	4,545
8	Rokiškio r.	1	14,2	4,559
9	Šalčininkų r.	1	13,6	3,308
10	Šilutės r.	1	11,8	3,122
11	Telšių r.	1	11	3,36
12	Ukmergės r.	1	15,2	4,625
13	Varėnos r.	1	14,2	3,381
14	Pakruojo r.	1	13,3	3,772
15	Šakių r.	1	14,1	4,4



7.7 pav. Miestų ir rajonų savivaldybių gyventojų mirtingumo funkcija

Šio tipo modeliai naudojami sociologijoje, psichologijoje, edukologijoje, rinkos tyrimuose. Tačiau daug dažniau naudojami regresijos modeliai, kuriuos sudaro ir kiekybiniai, ir kiekybiniai kintamieji.

7.2 pavyzdys. Vieno kiekybinio ir vieno kokybinio kintamojo regresija

Regresijos modeliai, kuriuos sudaro ir kiekybiniai, ir kokybiniai kintamieji dar vadinami ANCOVA (*Analysis-of-Covariance*) modeliais. Papildykime (7.19) modelį vienu kiekybiniu kintamuoju – sergamumas piktybiniais navikais, t. y. susirgusiųjų piktybiniais navikais 1000-iui gyventojų skaičius:

$$y_i = \alpha_1 + \alpha_2 k_i + \beta x_i + \varepsilon_i; \quad (7.21)$$

čia: y_i – vidutinis gyventojų mirtingumas,

x_i – tenka susirgusiųjų piktybiniais navikais skaičius 1000-iui gyventojų;

$$k_i = \begin{cases} 0, & \text{jei savivaldybė yra miesto,} \\ 1, & \text{jei savivaldybė yra rajono.} \end{cases}$$

Modelį (7.21) sudaro vienas nepriklausomas kiekybinis kintamasis (sergamumas piktybiniais navikais) ir vienas kokybinis kintamasis (dviejų lygių, kategorijų: miestai ir rajonai) – savivaldybės.

Kokia yra (7.21) formulės prasmė? Jei, kaip įprasta, atsitiktinės paklaidos ε_i vidurkį prilyginsime nuliui ($E(\varepsilon_i)=0$), tai: vidutinis gyventojų mirtingumas miestuose

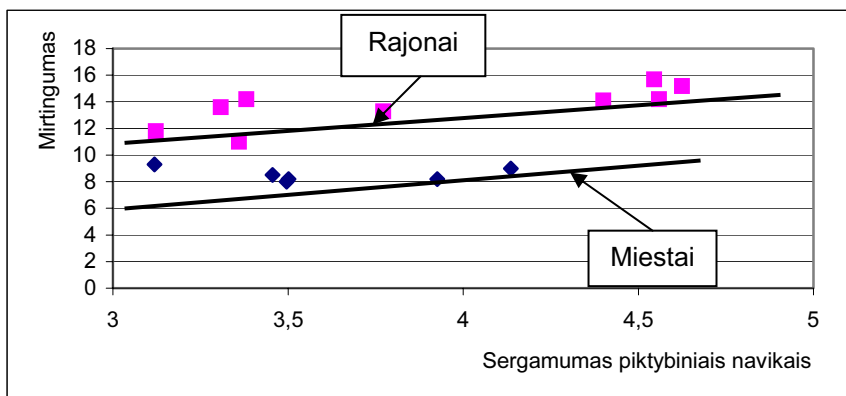
$$E(y_i | x_i, k_i = 0) = \alpha_1 + \beta x_i;$$

vidutinis gyventojų mirtingumas rajonuose

$$E(y_i | x_i, k_i = 1) = (\alpha_1 + \alpha_2) + \beta x_i,$$

Geometriškai ši situacija pavaizduota 7.8 pav. Modelis (7.21) grindžiamas prielaida, kad miestų ir rajonų gyventojų mirtingumo ir serga-

mumo piktybiniais navikais ryšio funkcija turi tą patį koeficientą β , tačiau skiriasi konstanta. Daroma prielaida, kad apskritai miestų gyventojų mirtingumo lygis skiriasi nuo rajonų gyventojų mirtingumo lygio, bet mirtingumo kitimas, priklausantis nuo sergamumo piktybiniais navikais, yra toks pat ir miestų, ir rajonų.



7.8 pav. Mirusiųjų skaičiaus ir sergamumo piktybiniais navikais miestuose ir rajonuose 1000-iui gyventojų sklaidos diagrama.

Hipotezė, kad koeficientas α_2 lygus nuliui (t. y., kad nėra jokio skirtumo tarp miestų ir rajonų gyventojų mirtingumo) gali būti patikrinta įvertinus (7.21) lygties koeficientus ir t kriterijumi patikrinus jų reikšmingumą. Pritaikę STATISTICA programą gautume, kad apskaičiuota koeficiento α_2 reikšmė $a_2=4,72$, o koeficiento β reikšmė $b=0,15$. Abu šie koeficientai yra statistiškai reikšmingi kai reikšmingumo lygmuo yra 5 proc. Vadinas, regresijos lygtis yra tokia:

$$\hat{y}_i = 3,29 + 4,72k_i + 1,45x_i. \quad (7.22)$$

Pastaba. Interpretuojant modelį, svarbu nesupainioti fiktyvaus kintamojo reikšmių ir atsiminti, kam buvo priskirta nulinė reikšmė. Grupė ar kategorija, kuriai ši reikšmė yra priskiriama, vadinama įvairiai – bazine, etalonine, lyginamąja. Interpretuojant ir lyginant remiamasi būtent šia

grupe. Nagrinėtuose modeliuose bazinė grupė (kategorija) yra miestų savivaldybės.

7.3 pavyzdys. Vieno kiekybinio kintamojo ir vieno daugiau nei dviejų kategorijų kokybinio kintamojo regresija

Tarkime, kad mus dominančios savivaldybių grupės yra trys: miestų savivaldybės, rajonų savivaldybės ir priemiestinių rajonų, esančių šalia didžiausių Lietuvos miestų, savivaldybės. 7.8 lentelę papildome šešiais rajonais, kuriems didelę įtaką daro didžiausi Lietuvos miestai (7.9 lentelė).

7.9 lentelė. Priemiestinių savivaldybių gyventojų mirtingumas ir sergamumas 1999 m.

Nr.	Savivaldybė	Kodas	Mirusiųjų 1000 gyv.	Susirgusių piktybiniais navikais 1000 gyv.
1	Kauno r.	priemiestinis	9,7	3,771
2	Klaipėdos r.	priemiestinis	12	3,831
3	Trakų r.	priemiestinis	11,3	3,807
4	Jonavos r.	priemiestinis	10,6	3,508
5	Vilniaus r.	priemiestinis	13,1	4,025
6	Šiaulių r.	priemiestinis	12,5	3,879

Šiuo atveju, kitaip negu anksčiau, nominalųjį kintamąjį sudaro trys kategorijos. Kaip sudaromas fiktyvus kintamasis, kai kategorijų daugiau nei dvi? Kai kategorijų yra daugiau nei dvi, sudaromas ne vienas, o keletas fiktyvių kintamųjų. ***Nustatant fiktyvių kintamųjų skaičių reikia vadovautis taisykle, kad fiktyvių kintamųjų turi būti vienu mažiau nei kokybinio kintamojo kategorijų.*** Pagal šią taisyklę, kai kategorijos yra trys, reikia sudaryti du fiktyvius kintamuosius. Regresijos modelis bus toks:

$$y_i = \alpha_1 + \alpha_2 k_{2i} + \alpha_3 k_{3i} + \beta x_i + \varepsilon_i; \quad (7.23)$$

čia: y_i – vidutinis gyventojų mirtingumas,
 x_i – susirgusių piktybiniais navikais 1000-iui gyventojų;

$$k_{2i} = \begin{cases} 1, & \text{jei savivaldybė yra rajono,} \\ 0, & \text{jei yra kitaip.} \end{cases}$$

$$k_{3i} = \begin{cases} 1, & \text{jei savivaldybė yra priemiestinė,} \\ 0, & \text{jei yra kitaip.} \end{cases}$$

Būtina atkreipti dėmesį, kad šiuo atveju reikšmė 0 yra priskirta miestų savivaldybėms. Jeigu mirtingumo skirtumo tarp miestų, priemiestinių ir rajonų savivaldybių nebūtų, regresijos lygtis būtų tokia:

$$E(y_i | k_{2i} = 0, k_{3i} = 0, x_i) = \alpha_1 + \beta x_i; \quad (7.24)$$

Tuo atveju, kai išsiskiria tik rajonų savivaldybės:

$$E(y_i | k_{2i} = 1, k_{3i} = 0, x_i) = (\alpha_1 + \alpha_2) + \beta x_i. \quad (7.25)$$

Jei nuo kitų skiriasi tik priemiestinės savivaldybės

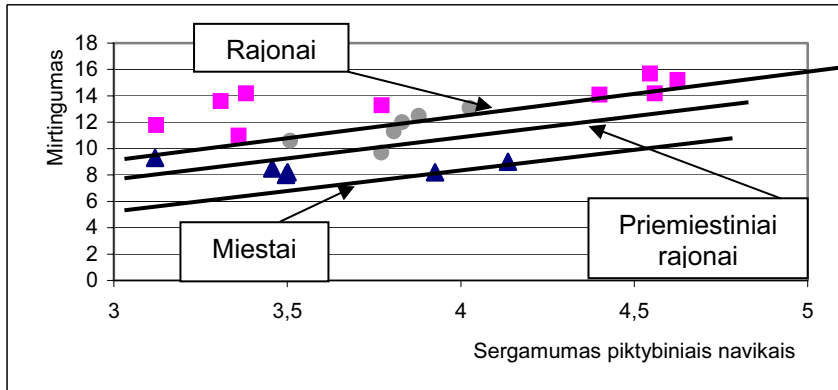
$$E(y_i | k_{2i} = 0, k_{3i} = 1, x_i) = (\alpha_1 + \alpha_3) + \beta x_i; \quad (7.26)$$

Apskaičiavę (7.23) lygties koeficientus gauname:

$$y_i = 2,79 + 4,68k_{2i} + 2,68k_{3i} + 1,59x_i. \quad (7.27)$$

Visi apskaičiuotieji koeficientai (išskyrus laisvąjį narį) yra statistiškai reikšmingi, kai reikšmingumo lygmuo yra 5 proc. (7.27) lygtį pavaizdavus grafiškai, diagrama būtų apytiksliai tokia, kokia yra 7.9 pav.

Gautų rezultatų interpretavimas priklauso nuo to, kaip sudaryti fiktyvūs kintamieji. Juos galima sudaryti įvairiais būdais. Išnagrinėtu atveju nulinė reikšmė priskirta miestų savivaldybėms, tačiau taip pat ją galima priskirti priemiestinių rajonų savivaldybėms arba rajonų savivaldybėms. Skirtingai sudarę fiktyvius kintamuosius, gausime kitokias (7.23) lygties koeficientų reikšmes, todėl, darant statistines išvadas, būtina labai atidžiai sekti, kaip sudaryti fiktyvūs kintamieji.



7.9 pav. Miestų, priemiestinių rajonų ir rajonų savivaldybių gyventojų sergamumo piktybiniais navikais ir mirusiųjų skaičiaus 1000-iai gyventojų sklaidos diagrama

Kai kokybinis kintamasis turi daug kategorijų, analizuoti šiuo metodu labai sudėtinga. Tokiu atveju geriau pasirinkti kitą statistinės duomenų analizės metodą.

7.8. Netiesinė (kreivinė) regresija

Daugelyje praktinių socialinių uždavinių tiesinis regresijos modelis, jei ir ne visai tiksliai, tai pakankamai gerai aprašo tikrąjį statistinio ryšio formą. Tačiau kartais pagal sklaidos diagramą arba remiantis žinomomis teorinėmis prielaidomis tenka pripažinti, kad ryšio forma yra aiškiai netiesinė. Netiesinės regresijos ir koreliacijos tema yra labai plati, todėl apsiribosime tik paprasčiausiomis netiesinių lygčių formomis. Laimė, šios nesudėtingos lygtys tinka daugeliui socialiniuose moksluose išskylančių uždavinių.

Polinominės regresijos modeliai

Bendriausias „netiesinis“ modelis yra *polinominė regresija*. Terminas „netiesinis“ parašytas kabutėse todėl, kad šis modelis iš esmės yra tiesinis. Tiksliau, šio modelio koeficientų skaičiavimo būdas niekuo ne-

siskiria nuo tiesinės regresijos koeficientų skaičiavimo. Pavyzdžiui: tar-kime, jūs matuojate asmenų psichologinį susijaudinimą jiems atliekant sudėtingą mažai žinomą užduotį. Remiantis psichologijoje gerai žinomu *Yerkes–Dodson* dėsnio, galima tikėtis, kad ryšys tarp susijaudinimo ir atlikimo bus netiesinis. Šį įsitikinimą galima užrašyti regresijos lygtimi:

$$\text{Atlikimas} = b_1 \cdot \text{Susijaudinimas} + b_2 \cdot \text{Susijaudinimas}^2 + \varepsilon, \quad \text{arba} \\ y_i = b_1 \cdot x_i + b_2 \cdot x_i^2 + \varepsilon_i;$$

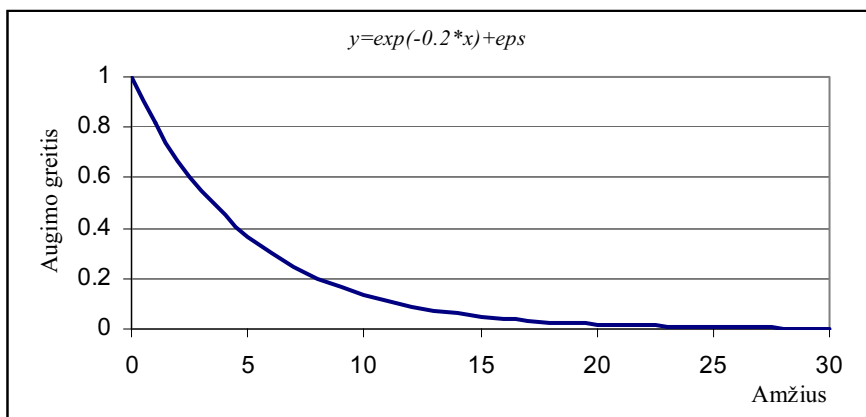
Šio modelio netiesiškumas yra išreikštas kintamuoju *Susijaudini-mas*², arba x_i^2 . Tačiau modelio koeficientai (b_1 ir b_2) nekeliami laipsniu, todėl modelis šių koeficientų atžvilgiu yra tiesinis.

Modeliai su netiesiniais parametrais

Išnagrinėkime pavyzdį apie ryšį tarp žmogaus amžiaus (kintamojo x) ir jo augimo greičio (kintamojo y). Akivaizdu, kad pirmaisiais gyve-nimo metais šių kintamųjų ryšys labai skirsis nuo suaugusio žmogaus. Yra žinoma, kad tokio tipo statistinį ryšį galima užrašyti eksponentine funkcija:

$$\text{Augimo greitis} = \exp(-\beta \cdot \text{Amžius}) + \varepsilon;$$

Žmogaus amžiaus ir jo augimo greičio sklaidos diagrama būtų maž-daug tokio pavidalo kreivė (7.10 pav.).



7.10 pav. Netiesinis ryšys tarp augimo greičio ir amžiaus

Atkreipkite dėmesį, kad šios lygties koeficientas yra rodiklyje. Tokio tipo modeliai – tai modeliai su netiesiniais parametrais.

Tačiau, nors šiuo atveju ryšys tarp x ir y yra netiesinis, jį labai lengvai galima padaryti tiesinį – apskaičiavus abiejų pusių logaritmus:

$$\ln(\text{Augimo greitis}) = -\beta \cdot \text{Amžius} + \varepsilon;$$

Dabar, pažymėję $\ln(\text{Augimo greitis})$ raide y , gausime standartinį tiesinės regresijos modelį be laisvojo nario (jis neįtrauktas dėl paprastumo). Norint apskaičiuoti tokio modelio koeficientus, pirmiausia ieškoma kintamojo *Augimo greitis* duomenų logaritmo, o po to – regresijos lygties koeficientų.

Ryšiai, kuriuos transformuojant galima gauti tiesinį pavidalą, yra „iš esmės tiesiniai“ (*angl. Intrinsically Linear Model*). Tokio tipo transformacijomis reikia naudotis visada, kada galima, kadangi šiuo atveju koeficientai apskaičiuojami tiesiogiai, o ne iteraciniais metodais.

Polinominė regresinė analizė Statistica programa

Atliekant polinominę regresinę analizę *Statistica* programa renka šios komandos:

Statistics

Advanced Linear/Nonlinear Models

General Linear Models

Polynomial regression

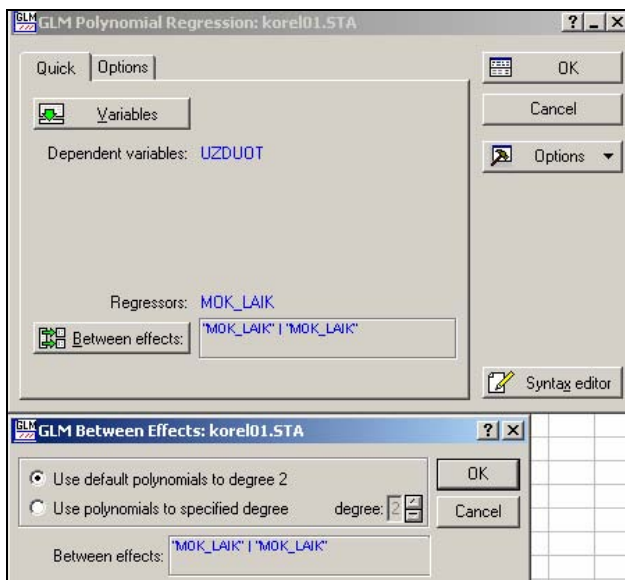
Dialoginiame lange (7.11 pav.) parenkami kintamieji, o klavišu „*Between effects*“ – polinomo laipsnis.

Šiuo skaičiavimo metodu gauti regresijos lygties koeficientai pateikiami rezultatų lange:

Summary

Coefficients.

Gautų rezultatų (7.10 lentelė) pateikimas ir analizė yra analogiška tiesinės regresijos modelio rezultatų analizei. Lygties koeficientai pateikti stulpelyje *Param*.



7.11 pav. Polinominės regresijos modelio kintamųjų ir polinomo eilės parinkimas

7.10 lentelė. Polinominės regresijos skaičiavimo rezultatai

<i>Parameter Estimates</i>								
<i>Effect</i>	<i>Param.</i>	<i>Std.Err</i>	<i>t</i>	<i>p</i>	<i>-95,00% Cnf.Lmt</i>	<i>+95,00% Cnf.Lmt</i>	<i>Beta (B)</i>	<i>St.Err.B</i>
<i>Intercept</i>	3,287	0,108	30,568	0,000	3,074	3,500		
MOK LAIK	2,510	0,145	17,284	0,000	2,222	2,799	4,078	0,236
MOK LAIK^2	-0,657	0,041	-16,201	0,000	-0,737	-0,576	-3,822	0,236

Užduotys

7.1. Remdamiesi Lietuvos nusikalstamumo prevencijos centro duomenimis (žr. 7.11 lentelę), ištyrkite bendro užregistruotų nusikaltimų skaičiaus ir šių nusikaltimų rūšių priklausomybę: a) sunkių nusikaltimų, b) tyčinių nužudymų, c) išžaginimų, d) vagysčių, e) automobilių vagys-

čių, f) plėšimų, g) turto prievartavimo, h) sukčiavimo, i) chuliganizmo. Sudarykite regresinį modelį, atlikite jo analizę ir paaiškinkite gautus rezultatus.

7.11 lentelė. Bendras užregistruotų nusikaltimų skaičius 10 tūkst. gyventojų ir nusikaltimų struktūrinės dalys

Metai	Nusikaltimų sk. 10 tūkst. gyv.	Sunkūs nusi- kaltimai	Tyčiniai nužudymai	Išžaginimai	Vagystės	Automobilių vagystės	Plėšimai	Turto prie- var- tavimas	Sukčiavimai	Chuliganiz- mas
1988	58,16		0,43	0,40						2,7
1989	84,16		0,43	0,40						2,7
1990	99,2	10,8	0,60	0,53	23572		0,90		1,2	3,0
1991	120,0	12,1	0,69	0,51	30486		1,07		1,0	3,2
1992	150,5	15,9	0,81	0,51	41053		1,30	0,12	1,2	3,8
1993	162,0	21,9	1,29	0,53	42291		1,98	0,63	2,2	3,9
1994	158,0	25,0	1,41	0,44	40252		11,33	0,85	2,4	4,6
1995	163,6	35,5	1,35	0,54	41619		7,64	0,61	2,7	6,9
1996	183,3	53,8	1,09	0,45	44600	10,5	9,38	0,65	2,9	8,1
1997	204,5	57,2	1,06	0,45	47193	10,3	10,72	0,64	3,6	8,4
1998	210,9	59,7	0,96	0,45	48213	11,2	9,85	0,52	3,9	8,0
1999	208,3	56,7	0,93	0,61	47865	9,9	9,13	0,62	2,5	7,3
2000	222,5	49,4	1,08	0,49	53225	14,0	11,82	0,52	1,8	7,3
2001	214,5	48,7	1,02	0,48	48899	15,8	11,33	0,53	2,3	6,5

7.2. 7.12 lentelėje pateikti Jungtinių Tautų duomenys apie įvairiose valstybėse užregistruotų nusikaltimų skaičių 100 000 gyventojų. Nustatykite, kaip pasikeitė valstybių diferenciacija pagal bendrą nusikalstamumo lygį šiuo laikotarpiu. Apskaičiuokite visus galimus rodiklius. Sudarykite diagramas ir suformuluokite išvadas.

7.12 lentelė. Bendras užregistruotų nusikaltimų skaičius 100 000 gyventojų

Metai	Baltarusija	Bulgarija	Kipras	Danija	Suomija	Graikija	Vengrija	Japonija	Lietuva
1980	471	506	1333	8282	10062	3063	1219	1178	410
1981	479	507	1365	8313	10596	3179	1260	1244	415
1982	456	526	1277	8512	10881	3424	1308	1291	435
1983	557	580	1496	9015	12040	3996	1421	1292	565
1984	588	552	1569	9756	13416	3562	1478	1324	585
1985	584	541	1481	10391	13390	2933	1567	1331	600
1986	541	624	1939	11091	14208	2954	1737	1302	573
1990	738	772	541	11531	8779	3256	3291	1397	992
1991	792	1877	522	11471	7812	3503	4256	1451	1200
1992	937	2387	539	11907	7792	3678	4332	1469	1505
1993	998	2328	599	12084	7644	3454	3895	1512	1620
1994	1161	2363	596	11878	7650	2909	3795	1491	1580
1995	1276	2465	550	10309	7472	3147	4907	1421	1636
1996	1236	2348	612	10043	7352	3336	4572	1409	1833
1997	1251	2909	523	10081	7273	3591	5066	1507	2045

7.3. Remdamiesi Lietuvos nusikalstamumo prevencijos centro 1990–2001 metų duomenimis, atlikite užregistruotų nusikaltimų Baltijos valstybėse regresinę analizę. Išstirkite užregistruotų nusikaltimų Baltijos valstybėse priklausomybę (žr. 7.13 lentelę), sudarykite prognozavimo modelį ir paaiškinkite gautus rezultatus.

7.13 lentelė. Užregistruoti nusikaltimai Baltijos valstybėse 10 tūkst. gyventojų

Metai	Lietuva	Latvija	Estija
1990	99,2	129,0	150,0
1991	120,0	151,0	206,0
1992	150,5	235,0	267,0
1993	162,0	204,0	245,0
1994	158,0	161,0	236,0
1995	163,6	155,0	265,0
1996	183,3	152,0	240,0
1997	204,5	147,0	280,0
1998	210,9	147,9	314,4
1999	208,3	180,2	356,7
2000	222,5	211,3	419,8
2001	214,5	215	428,0

7.4. Atlikite asmenų, kaltinamų padarius nusikaltimus (žr. 7.14 lentelę) socialinės bei ekonominės struktūros dinamikos analizę. Gavę skaičiavimų rezultatus, suformuluokite išvadas. Patikrinkite šių darbinių hipotezių teisingumą.

Tarp asmenų, kaltinamų padarius nusikaltimus, kasmet daugėja:

- moterų;
- nepilnamečių;
- neturinčių nuolatinio darbo;
- pakartotinai nusikaltusių asmenų;
- neblaivių asmenų.

7.14 lentelė. Asmenų, kaltinamų padarius nusikaltimus, struktūra

Metai	Visi kaltinami asmenys (10 tūkst. gyv.)	Moterys (proc.)	Nepilnamečiai (proc.)	Nedirbo ir nesimokė (proc.)	Nusikalto pakartotinai (proc.)	Neblaivūs (proc.)
1988	34,9		11,9			
1989	34,8		14,6			
1990	33,7	12,5	16,3	22,4	20,2	37,9
1991	35,5	10,1	14,5	28,9	18,9	39,8
1992	50,3	10,2	14,6	38,8	18,5	35,9
1993	54,8	10,3	15,6	53,4	21,3	37,4
1994	57,2	11,4	14,3	60,4	24,0	38,1
1995	61,8	10,4	14,7	61,7	21,7	33,9
1996	60,0	11,4	15,3	63,2	23,2	32,9
1997	68,9	13,9	13	63,2	22,6	29,2
1998	68,5	13,5	13,1	62,7	22,7	27,6
1999	68,0	13,0	13,3	62,7	23,4	25,7
2000	67,7	10,0	14,3	66,5	27,4	28,1
2001	72,2	8,7	13,8	67,2	29,6	30,8

7.5. Remdamiesi statistiniais duomenimis (žr. 7.15 lentelę) apie kai kurių turtingiausių pasaulio valstybių išsivystymo lygį 1991 metais (*New York: Bantam books, 1992; <http://www.korpios.org>*), raskite statistiškai reikšmingus ryšius tarp socialinių bei ekonominių ir kriminalinių rodiklių. Pateikite išvadas.

7.15 lentelė. Socialiniai bei ekonominiai ir kriminaliniai kai kurių turtingiausių pasaulio valstybių rodikliai

Valstybė	Nužudymų (100 tūkst. gyv.)	Išžaginių (100 tūkst. gyv.)	Kalinių (1000 gyv.)	Vidurinioji klasė (proc.)	Perkamoji galia (JAV dol.)	Nesantuokoje gimusių vaikų (proc.)
Kanada	5,45	–	–	58,2	19178	12,1
Danija	5,17	11,23	0,7	–	17621	41,9
Vokietija	4,20	8,60	0,8	70,1	19500	9,4
Norvegija	1,99	7,87	–	73,0	16904	–
Jungtinė karalystė	1,97	7,26	1,0	58,5	–	19,2
Švedija	1,73	15,70	0,6	79,0	19729	46,4
Japonija	1,20	1,40	0,4	90,0	19107	1,0
Suomija	0,70	7,20	–	–	15997	–
JAV	8,40	37,20	4,2	53,7	22204	21,5



8. POLITINIO GYVENIMO STATISTIKA

„Patarimai retai kada priimami su dėkingumu: kam jie labiausiai reikalingi, rečiausiai jais naudojasi.“

Lordas Česterfildas

Politinio gyvenimo statistika nagrinėja valstybės politinėje sistemoje vykstančių procesų kiekybines charakteristikas. Statistiniai duomenys ir jų analizė padeda atskleisti valstybėje vykstančius kitimus, jos raidos dėsningumus, pamatyti ir įvertinti naujus reiškinius. Politinio gyvenimo statistika padeda suvokti ir apibūdinti valstybės valdymo institucijų poreikius, tiksliau nustatyti ekonominės ir socialinės politikos gaires, numatyti politinės sistemos raidos kelius.

Ypač daug politinės statistikos gyventojams pateikiama prasidėjus rinkimų kampanijoms. Per visuomenės informavimo priemones pateikiami gyventojų nuomonių apklausų rezultatai, rinkėjų nusiteikimo dalįvauti rinkimuose analizė, politikų reitingų ir populiarumo suvestinės. Ši informacija pateikiama skaičiais, indeksais, lentelėmis, diagramomis. Po rinkimų gautus rezultatus bandoma analizuoti ir komentuoti šiuo tikslu taikant sudėtingus statistinės analizės metodus. Žiūrėdami į pateikiamus statistinių tyrimų rezultatus, daugelis būsimų rinkėjų ir žurnalistų dažnai gerai ir nesupranta, kad diagramose ar lentelėse pateikiami skaičiai (dažniausiai procentais) nėra tikslūs – tai tik įverčiai. Nesuprantant gauto rezultato patikimumo sunku suvokti, kad parametrų skirtumas, gautas iš nedidelės ar nereprezentatyvios imties, iš tiesų gali būti nereikšmingas.

Visos politinės partijos domisi ir prognozuoja atskirų visuomenės sluoksnių elgesį. Jei imtis yra sudaryta taip, kad ji reprezentuoja bendrą populiaciją, tai ji nebūtinai yra tinkama, norint įvertinti šios populiacijos dalių savybes. Tam tikros partijos rėmėjai, įtraukti į tokią imtį, nebūtinai reprezentuoja visus partijos rėmėjus, o, pavyzdžiui, 18–25 metų jaunuoliai, patekę į imtį, nebūtinai reprezentuoja visus 18–25 metų jaunuolius.

Norint gauti tikslus ir pagrįstus rezultatus, neužtenka taikyti statistikos ir duomenų analizės metodus. Būtina kartu gerai išmanyti tiriamą objektą ir turėti specialių tiriamos srities žinių.

8.1. Rinkimų prognozės

Rinkimų prognozės, taip pat ir kiti politinės statistikos metodai jau seniai ir plačiai taikomi vakarų Europos ir JAV šalyse. Pavyzdžiui, Amerikos NES (Nacionalinių rinkimų tyrimai) jau nuo 1952 m. apklausia rinkėjų imtį 2 mėn. prieš rinkimus ir organizuoja porinkiminę apklausą su kiek įmanoma tais pačiais respondентаis. Statistinių metodų išplitimas ir taikymas priklauso nuo kiekvienoje šalyje esančių sąlygų – politikos istorijos, rinkimų sistemos, rinkimų organizavimo būdų ir pan. Politinės statistikos tyrimus galima skirti į dvi pagrindines grupes:

- atrankinių grupių nuomonių apklausa;
- panelinis tyrimas

Atrankinių grupių nuomonių apklausos tyrimo tikslas yra gauti žinių apie rinkėjų visumą. Kaip ir kiekvienoje nuomonių apklausoje, pradinė problema yra reprezentatyvios imties atrinkimas. Atliekant nuomonių apklausą kelis kartus, apklausai atrenkamos skirtingos imtys.

Panelinio tyrimo modelis pastaruoju metu taikomas vis plačiau. Jis konstruojamas taip, kad tie patys asmenys būtų apklausiami du ar daugiau kartų, kad būtų galima stebėti jų nuomonių apie tiriamą procesą ar reiškinį pokyčius. Toks metodas tiksliau ir aiškiau parodo vykstančius visuomenėje pokyčius nei atskirais laikotarpiais vykdoma skirtingų respondentų nuomonių apklausa.

Sudarant imtį būtina įvertinti tiek bendrus faktorius (amžių, socialinę bei finansinę padėtį ir pan.), tiek ir ankstesnių rinkimų rezultatus bei rinkėjų elgesį ankstesnių rinkimų metu. Lietuvoje dažnai taikomas geografinis metodas, t.y. imtys sudaromos pagal gyvenamąją vietą, mažai skiriant dėmesio kitoms charakteristikoms. Žinoma, kai elektoratas yra homogeninis t. y., rinkėjų požūriai yra panašūs, tai nesvarbu, kaip bus sudaryta imtis, vis tiek apklausos rezultatai bus panašūs. Jei ankstesnių rinkimų rezultatai rodo, kad rinkėjų populiacija nėra vienalytė, tai sudarant imtį šią informaciją būtina įvertinti.

Priešrinkiminėse apklausose aktuali išlieka atsisakymo dalyvauti apklausose problema, ir tuo galima iš dalies paaiškinti rezultatų neįver-

tinimą ar perversinį. Įprasta, kad apklausoje atsisako dalyvauti 15–20 proc. respondentų. Tai dažniausiai neįvertinama analizuojant apklausos rezultatus. Vadinasi, išdalijant 1000 anketų iš tiesų tenka bendrauti su 1200 respondentų, iš kurių 200 atsisakius dalyvauti apklausoje, rezultatai vis vien vertinami taip, tarsi visi būtų sutikę dalyvauti. Kita problema – neapsisprendę rinkėjai, kurie apsunkina rinkimų rezultatų prognozę.

Rinkėjų dalyvavimas rinkimuose

Rinkėjų dalyvavimas rinkimuose yra aktuali politinio mokslo tyrimo sritis. Aukštas rinkėjų dalyvavimo lygis turi didelės reikšmės visai politinei sistemai, taip pat turi įtakos rinkimų procesui, kandidato išrinkimui ir jo atsakomybei. Deja, rinkėjų dalyvavimo rinkimuose statistika nėra džiuginanti. Pavyzdžiui, 1996 m. Lietuvos Respublikos Seimo rinkimuose dalyvavo 1374673, arba 52,92 proc. rinkėjų; 2000 m. rinkimuose dalyvavo 1539743, arba 58,63 proc. rinkėjų. 1998 m. JAV Kongreso rinkimuose dalyvavo tik 45,3 proc. balsavimo teisę turinčių piliečių. Balsavimas rinkimuose yra apibūdinamas kaip svarbiausias demokratijos laimėjimas ir yra viena iš pagrindinių piliečių teisių, taip pat tai parodo ir piliečių atsakomybę.

Galima išskirti keletą būdingų viso demokratinio pasaulio rinkėjų elgesio savybių:

vyresnio amžiaus žmonės balsuoja dažniau nei jaunesnio;

labiau išsilavinę balsuoja dažniau nei mažiau išsilavinę;

žmonės, turintys didesnes pajamas, balsuoja dažniau, nei turintys mažas pajamas;

žmonės, gyvenantys nuosavuose namuose ir butuose, balsuoja dažniau, nei nuomininkai.

Kai kuriose valstybėse (Belgijoje, Argentinoje, Australijoje, Graikijoje, Liuksemburge ir Singapure) balsavimas yra privalomas; jis yra reguliuojamas rinkimų įstatymais. 1892 m. Belgija įvedė priverstinę rinkimų sistemą. Nedalyvavimas rinkimuose gali turėti tam tikrų pasekmių. Pavyzdžiui: Belgijoje nedalyvaujantiems rinkimuose gali būti sudėtinga gauti darbą viešajame sektoriuje. Graikijoje tai gali apsunkinti naujo paso ar vairuotojo pažymėjimo gavimą. Skirtumas tarp rinkėjų aktyvumo pagal laisvą ir priverstinę balsavimo sistemą yra didelis. Palyginkime dvi valstybes su skirtingomis rinkimų sistemomis – Kanadą ir Australiją (8.1 lentelė).

8.1 lentelė. Rinkimuose dalyvavę rinkėjai (proc.)

Metai	Kanados federalinės vyriausybės rinkimai (proc.)	Metai	Australijos Senato rinkimai (proc.)
2000	61	2001	95,19
1997	67	1998	95,34
1993	70	1996	96,20
1988	76	1993	96,22

Šaltiniai: Elections Canada and Elections Alberta;
AEC (Australian Electoral Commission).

Dėl kokių priežasčių rinkėjai nustoja domėtis rinkimais? Kaip teigia Kanados tyrimų ir informacijos centras (*Centre for Research and Information on Canada*), pagrindinės priežastys, kurias nurodo rinkėjai, yra šios:

- 86 proc. kanadiečių populiacijos tiki, kad politikai meluoja norėdami būti išrinkti;
- 13 proc. kanadiečių populiacijos gerbia ir pasitiki politinėmis partijomis;
- 87 proc. kanadiečių populiacijos mano, kad vyriausybė turėtų labiau konsultuotis su gyventojais;
- 28 proc. kanadiečių populiacijos mano, kad vidutinis pilietis turi įtakos vyriausybei;
- 79 proc. kanadiečių populiacijos mano, kad vidutinis gyventojas turėtų turėti įtakos vyriausybei.

Politinė statistika labiau decentralizuota nei kitos statistikos rūšys. Valstybinės statistikos įstaigos paprastai politinės statistikos duomenų nerenka. Rinkimų duomenis pateikia Vyriausioji rinkimų komisija, sociologines priešrinkimines apklausas partijų ir atskirų kandidatų užsakymu atlieka valstybinės ir privačios sociologinės tarnybos. Šios apklausos skiriasi ir savo apimtimi, ir sudėtingumu. Tokios apklausos pavyzdys – anketa, kurią prieš 2002 m. prezidento rinkimus savo tyrimuose naudojo viešosios nuomonės ir rinkos tyrimų kompanija „TNS GALLUP“.

ŠIOJE ŠEIMOJE APKLAUSIAMA (-AS) _____
 (ĮRAŠYKITE RESPONDENTO, KURIS BUS APKLAUSIAMAS, LYTĮ.

TOLESNEI APKLAUSAI PASIRINKITE JUMS REIKIAMOS LYTIES ŽMOGŲ, KURIO GIMTADIENIS TARP TOS PAČIOS LYTIES ŠEIMOS NARIŲ YRA ARČIAUSIAI NUO JŪSŲ LANKYMOŠI ŠIAME BUTE DATOS.

A 1. Ar jūs turite teisę balsuoti?

- | | |
|---------|-------------------------|
| 1. Taip | 1 |
| 2. Ne | 2 ⇒ NUTRAUKITE APKLAUSĄ |

NORĖTUME JUMS UŽDUOTI KELETĄ KLAUSIMŲ APIE LIETUVOS POLITINĮ GYVENIMĄ IR ARTĖJANČIUS PREZIDENTO RINKIMUS.

K1. Jei kalbėtume apie gyvenimą Lietuvoje, ar, Jūsų manymu, apskritai jis gerėja, blogėja, ar nekinta?

- | | |
|-------------------------|---|
| Gyvenimas gerėja | 1 |
| Nei gerėja, nei blogėja | 2 |
| Gyvenimas blogėja | 3 |
| Nežinau | 4 |

K2. Kaip vertintumėte dabartinės vyriausybės veiklą? Kaip pasakytumėte: ji vykdo savo pareigas labai gerai, gerai, blogai ar labai blogai?

- | | |
|--------------|---|
| Labai gerai | 1 |
| Gerai | 2 |
| Blogai | 3 |
| Labai blogai | 4 |
| Nežinau | 5 |

K3. Kaip vertintumėte dabartinio Prezidento veiklą? Kaip pasakytumėte: jis vykdo savo pareigas labai gerai, gerai, blogai ar labai blogai?

- | | |
|-------------|---|
| Labai gerai | 1 |
| Gerai | 2 |

Blogai	3
Labai blogai	4
Nežinau	5

K3.a. Kaip Jūs manote, ar Lietuvos Respublikos Prezidentui turėtų būti suteikta daugiau vykdomosios valdžios teisių?

Taip, turėtų	1
Ne, neturėtų	2
Nežinau	3

K4. Pagalvokite, ką valdžia visų pirmiausia turėtų padaryti, kad gyvenimas Lietuvoje taptų geresnis. Tarkime, kad Jums reikėtų pasirinkti tris dalykus, kuriuos valdžia turėtų nuveikti, kas tai būtų.

NEPASAKINĖKITE RESPONDENTUI. ĮRAŠYKITE RESPONDENTO ATSAKYMUS:

1. Pirma _____
2. Antra _____
3. Trečia _____

K5. Pasakykite, kurios iš šių problemų, Jūsų manymu, yra pačios aktualiausios Lietuvai šiuo metu. Ką visų pirmiausia šiuo metu turėtų nuveikti Lietuvos valdžia? Pasirinkite ne daugiau kaip septynias problemas.

Viename stulpelyje prašome žymėti tik po vieną atsakymo variantą.

	1	2	3	4	5	6	7
1. Atstatyti Valdovų rūmus Vilniuje	1	1	1	1	1	1	1
2. Daugiau finansiškai remti žemės ūkį	2	2	2	2	2	2	2
3. Didinti išmokas socialiai remtiniams žmonėms (neįgaliesiems, pensininkams, daugiavaikėms šeimoms, vienišoms motinoms ir pan.)	3	3	3	3	3	3	3
4. Gražinti bankrutavusių bankų indėlius, bankrutavusių įmonių skolas darbuotojams	4	4	4	4	4	4	4

5. Gražinti nuosavybę buvusiems savininkams	5	5	5	5	5	5	5
6. Griežčiau ginti darbuotojų teises (kad būtų laiku mokamas atlyginimas, suteikiamos priklausančios atostogos, mokama už viršvalandžius ir pan.)	6	6	6	6	6	6	6
7. Ieškoti rinkų parduoti žemės ūkio produkciją	7	7	7	7	7	7	7
8. Įvesti alternatyvią SODRAI pensijų sistemą	8	8	8	8	8	8	8
9. Kovoti su girtuoklyste	9	9	9	9	9	9	9
10. Kovoti su narkomanijos plitimu	10	10	10	10	10	10	10
11. Kovoti su valstybės pareigūnų korupcija	11	11	11	11	11	11	11
12. Kuo greičiau įstoti į Europos Sąjungą	12	12	12	12	12	12	12
13. Kuo greičiau įstoti į NATO	13	13	13	13	13	13	13
14. Mažinti valdžios kišimąsi į žmonių gyvenimą	14	14	14	14	14	14	14
15. Neleisti imigracijos į Lietuvą	15	15	15	15	15	15	15
16. Sugriežtinti bausmes už finansinius nusikaltimus	16	16	16	16	16	16	16
17. Padidinti minimalų darbo užmokestį	17	17	17	17	17	17	17
18. Padidinti mokesčius turtingiesiems	18	18	18	18	18	18	18
19. Padidinti pensijas	19	19	19	19	19	19	19
20. Panaikinti karinę prievolę	20	20	20	20	20	20	20
21. Pasirūpinti, kad Lietuvoje būtų mažiau pažeidžiamos žmogaus teisės	21	21	21	21	21	21	21
22. Remti jaunimą (išsimokslinimas, darbo vietos, lengvatinės paskolos ir pan.)	22	22	22	22	22	22	22
23. Siekti, kad nebūtų prieš laiką uždaryta Ignalinos atominė elektrinė	23	23	23	23	23	23	23
24. Siekti, kad sumažėtų komunaliniai mokesčiai	24	24	24	24	24	24	24
25. Skatinti krikščionišką moralę	25	25	25	25	25	25	25
26. Skatinti naujų darbo vietų steigimą	26	26	26	26	26	26	26
27. Skirti didesnę finansavimą mokslui ir švietimui	27	27	27	27	27	27	27
28. Sudaryti išskirtines sąlygas smulkiam ir vidutiniam verslui	28	28	28	28	28	28	28
29. Sudaryti išskirtines sąlygas užsienio investicijoms Lietuvoje	29	29	29	29	29	29	29

30. Sugriežtinti bausmes už kriminalinius nusikaltimus	30	30	30	30	30	30	30
31. Sugriežtinti mokesčių surinkimą	31	31	31	31	31	31	31
32. Sumažinti mokesčius	32	32	32	32	32	32	32
33. Sustabdyti prekių kontrabandą	33	33	33	33	33	33	33
34. Suteikti būstą neišgalintiems jo nusi-pirkti	34	34	34	34	34	34	34
35. Uždaryti Būtingės naftos terminalą, kad būtų išsaugotas Lietuvos pajūris	35	35	35	35	35	35	35
36. Užtikrinti tautinėms mažumoms gali-mybę mokytis savo gimtąja kalba	36	36	36	36	36	36	36

Įsivaizduokite, kad Jums reikia paskirstyti valstybės biudžetą.

K6.1. Kurioms sritims Jūs skirtumėte daugiau lėšų, nei yra skiriama šiuo metu? PASIRINKITE NE DAUGIAU NEI TRIS SRITIS.

K6.2. O kurioms sritims Jūs skirtumėte mažiau lėšų, nei yra skiriama šiuo metu? PASIRINKITE NE DAUGIAU NEI TRIS SRITIS.

	Daugiau lėšų	Mažiau lėšų
1. Sveikatos apsaugai, vaistų kompensavimui, ligoninėms	1	1
2. Krašto apsaugai, kariuomenei	2	2
3. Švietimui, mokykloms, universitetams	3	3
4. Pensijoms, socialinėms pašalpoms	4	4
5. Policijai ir kitoms teisėsaugos struktūroms	5	5
6. Muitinei	6	6
7. Kultūrai, menui	7	7
8. Sportui	8	8
9. Žemės ūkiui	9	9
10. Valstybės valdymo reikmėms	10	10
11. Aplinkos apsaugai	11	11
12. Kita (parašykite)	12	12

K7. Kalbant apie būsimus Prezidento rinkimus gruodžio mėnesį, ar Jūs tikrai balsuosite, greičiausiai balsuosite, gal balsuosite, bet gal ir ne, ar tvirtai nusprendėte nebalsuoti?

Tikrai balsuosiu	1	
Greičiausiai balsuosiu	2	
Galbūt balsuosiu, bet galbūt ir ne	3	
Greičiausiai nebalsuosiu	4	
Tvirtai nusprendžiau nebalsuoti	5	⇒ PEREIKITE PRIE KLAUSIMO K12
Nežinau	6	

PADUOKITE RESPONDENTUI KORETELEJ SU KANDIDATŲ PAVARDĖMIS

K8.1. Už kurį iš išvardintų kandidatų Jūs ketinate balsuoti per ateinančius Prezidento rinkimus gruodžio mėnesį?

JEI RESPONDENTAS PASAKĖ „NEŽINAU“ ARBA „NĖ UŽ VIENĄ“, KLAUSTI KLAUSIMĄ K8.4.

K8.2. O už kurį kandidatą iš likusiųjų Jūs balsuotumėte, jei Jūsų pasirinktas kandidatas nedalyvautų rinkimuose?

K8.3. O kuris kandidatas būtų trečias Jūsų eilėje?

K8.4. O dabar prašau pagalvoti ir pasakyti, už kuriuos kandidatus Jūs tikrai nebalsuotumėte?

K8.5. KLAUSTI TIK TŲ, KURIE KLAUSIME K8.1. PASIRINKO ATSAKYMO VARIANTĄ „NEŽINAU“. KITU ATVEJU PEREITI PRIE KLAUSIMO K9.

K8.5. Nors sakote, kad dar nežinote, už ką balsuosite Prezidento rinkimuose, tačiau visgi kuris iš kandidatų yra Jums artimiausias? PEREITI PRIE KLAUSIMO K10.

	K8.1. Pir- mas pasi- rinkimas	K8.2. Antras pasirinkimas	K8.3. Tre- čias pasi- rinkimas	K8.4. Ne- balsuotų	K8.5. Ar- timiausias
1. A. Brazauskas	1	1	1	1	1
2. A. Kubilius	2	2	2	2	2
3. A. Paulauskas	3	3	3	3	3
4. Č. Juršėnas	4	4	4	4	4
5. E. Gentvilas	5	5	5	5	5
6. K. Bobelis	6	6	6	6	6

7. K. Glaveckas	7	7	7	7	7
8. K. Prunskienė	8	8	8	8	8
9. R. Paksas	9	9	9	9	9
10. V. Adamkus	10	10	10	10	10
11. V. Andriukaitis	11	11	11	11	11
12. V. Šustauskas	12	12	12	12	12
13. Kitas	13	13	13	13	13
14. Nė už vieną	14	14	14	14	14
15. Nežinau	15	15	15	15	15

KLAUSIMĄ K9 KLAUSTI TIK TUOS, KURIE K8.1. PASIRINKO KURĮ NORS IŠ KORTELĖJE ESANČIŲ KANDIDATŲ.

K9. Jūs sakėte, kad pirmiausia Jūs balsuotumėte už (PERSKAITYKITE KANDIDATO, KURĮ RESPONDENTAS PASIRINKO KLAUSIME K8.1 PAVARDE). Ar esate visiškai tikras šiuo savo pasirinkimu?

Taip, visiškai tikras	1
Daugmaž tikras	2
Netikras	3
Sunku pasakyti	4

K10. Norėčiau, kad kiekvieną iš galimų kandidatų į Prezidentus trumpai apibūdintumėte. Pasakykite, kokios, Jūsų manymu, geros ar blogos savybės charakterizuoja šį žmogų.

PERSKAITYKITE KANDIDATŲ PAVARDES PO VIENĄ IR UŽRAŠYKITE RESPONDENTŲ ATSAKYMUS.
NEPASAKINĖKITE.

1. A. Brazauskas
2. A. Kubilius
3. A. Paulauskas
4. Č. Juršėnas
5. E. Gentvilas
6. K. Bobelis
7. K. Glaveckas
8. K. Prunskienė
9. R. Paksas

14. Moka bendrauti su žmonėmis	14	14	14	14	14	14	14	14	14	14	14	14
15. Moka kalbėti	15	15	15	15	15	15	15	15	15	15	15	15
16. Naujoviškas	16	16	16	16	16	16	16	16	16	16	16	16
17. Nepriekaištingos praeities	17	17	17	17	17	17	17	17	17	17	17	17
18. Nepriklauso nuo jokios partijos	18	18	18	18	18	18	18	18	18	18	18	18
19. Neryžtingas	19	19	19	19	19	19	19	19	19	19	19	19
20. Paprastas, artimas žmonėms	20	20	20	20	20	20	20	20	20	20	20	20
21. Pasipūtęs / arogantiškas	21	21	21	21	21	21	21	21	21	21	21	21
22. Pasišventęs valstybės labui	22	22	22	22	22	22	22	22	22	22	22	22
23. Patrauklus	23	23	23	23	23	23	23	23	23	23	23	23
24. Per jaunas	24	24	24	24	24	24	24	24	24	24	24	24
25. Per senas	25	25	25	25	25	25	25	25	25	25	25	25
26. Principingas, turintis savo nuomonę	26	26	26	26	26	26	26	26	26	26	26	26
27. Protingas	27	27	27	27	27	27	27	27	27	27	27	27
28. Rūpinasi paprastais žmonėmis	28	28	28	28	28	28	28	28	28	28	28	28
29. Solidus	29	29	29	29	29	29	29	29	29	29	29	29
30. Sugeba vadovauti	30	30	30	30	30	30	30	30	30	30	30	30
31. Susikompromitavęs	31	31	31	31	31	31	31	31	31	31	31	31
32. Tolerantiškas	32	32	32	32	32	32	32	32	32	32	32	32
33. Turi gerą humoro jausmą	33	33	33	33	33	33	33	33	33	33	33	33
34. Turi išmanančią ir gabią padėjėjų komandą	34	34	34	34	34	34	34	34	34	34	34	34
35. Turi tvirtą užnugarį	35	35	35	35	35	35	35	35	35	35	35	35

Dabar norėtume užduoti dar keletą klausimų apie jus.

K12. Kaip Jūs manote, ar esate kurios nors politinės partijos šalininkas?

Esu stiprus, įsitikinęs vienos iš politinių partijų šalininkas 1

Esu vienos iš politinių partijų šalininkas, bet nelabai stiprus 2

Nesu jokios politinės partijos šalininkas 3

⇒ Pereiti prie klausimo K14

Nežinau 4

⇒ Pereiti prie klausimo K14

PADUOKITE KORTELĘ NR. 4 SU PARTIJŲ SĄRAŠU.

K13. Kurios politinės partijos šalininkas Jūs esate?

Lietuvos centro sąjungos	1
Lietuvos krikščionių demokratų	2
Lietuvos laisvės sąjungos	3
Lietuvos lenkų rinkimų akcijos	4
Lietuvos liberalų demokratų partijos	5
Lietuvos liberalų sąjungos	6
Lietuvos socialdemokratų partijos	7
Moderniųjų krikščionių demokratų sąjungos	8
Naujosios sąjungos (socialliberalų)	9
Tėvynės sąjungos (Lietuvos konservatorių)	10
Valstiečių ir Naujosios demokratijos sąjungos	11
Kitos partijos	12
Nežinau / esu neapsisprendęs	13

K14. Nepaisant to, kurios partijos šalininkas Jūs esate, už kurią partiją balsuotumėte rinkimuose į LR Seimą, jei artimiausią sekmadienį vyktų balsavimas?

Lietuvos centro sąjungą (K. Glaveckas)	1
Lietuvos krikščionių demokratų (K. Bobelis, A. Saudargas)	2
Lietuvos laisvės sąjungą (V. Šustauskas)	3
Lietuvos lenkų rinkimų akciją (G. Mincevič, V. Tomaševski)	4
Lietuvos liberalų demokratų partiją (R. Paksas)	5
Lietuvos liberalų sąjungą (E. Gentvilas)	6
Lietuvos socialdemokratų partiją (A. Brazauskas)	7
Moderniųjų krikščionių demokratų sąjungą (V. Bogušis)	8
Naująją sąjungą (socialliberalus) (A. Paulauskas)	9
Tėvynės sąjungą (Lietuvos konservatorius) (V. Landsbergis, A. Kubilius)	10
Valstiečių ir Naujosios demokratijos sąjungą (K. Prunskienė, R. Karbauskis)	11
Kitą partiją	12
Nebalsuočiau	13
Nežinau / esu neapsisprendęs	14

K15. Kai kurie politikai tvirtina, kad dešinėsios partijos turėtų kelti vieną bendrą kandidatą į Prezidentus. Ar Jūs pritariate šiai idėjai?

1. Taip, pritariu
2. Pritariu, jei tas kandidatas būtų (įrašyti)
3. Ne, nepritariu
4. Neturiu nuomonės

K16. Ar dažnai pastaruoju metu Jūs einate į bažnyčią arba religinius susirinkimus, neskaitant vestuvių, laidotuvių ir krikštynų?

GALIMAS TIK VIENAS ATSAKYMAS.

Dažniau nei kartą per savaitę	1
Kartą per savaitę	2
Kartą per mėnesį	3
Tik per didžiąsias religines šventes	4
Kartą per metus	5
Rečiau	6
Beveik niekada, niekada	7

K17. Nepriklausomai nuo to, ar Jūs einate į bažnyčią, ar ne, kaip Jūs save apibūdintumėte?

GALIMAS TIK VIENAS ATSAKYMAS.

Giliai tikintis	1
Tikintis	2
Abejojantis	3
Netikintis	4
Nežinau	5

Klausimai interviuotojui

Interviu (apklausos) data _____ mėn. _____ d.

Interviu (apklausos) pradžia _____ val. _____ min.

Interviu (apklausos) trukmė _____ min.

Kelintas tai buvo Jūsų interviu šiandien? _____ int.

**X2. Ar patalpoje, kurioje vyko interviu, be žmogaus, kurį apklausi-
nėjote, buvo (ar užėjo kelioms minutėms) dar koks nors asmuo?**

Pažymėkite visus tinkamus variantus

1. Nieko nebuvo, neužėjo
2. Vaikai iki 6m. amžiaus
3. 6 m. amžiaus vaikai ir vyresni
4. Sutuoktinis(-ė)
5. Kiti giminaičiai
6. Kiti suaugusieji (ne giminės)

8.2. Politinio gyvenimo statistikos metodai

Remiantis US Census Bureau tyrimais (*Population Division Working Paper No 44*), galima sakyti, kad rinkimų statistikos analizė turėtų atsakyti į keletą pagrindinių klausimų:

- kokie socialiniai ir demografiniai rodikliai turi įtakos rinkėjų nuomonių pokyčiams;
- kaip rinkėjų socialiniai ir demografiniai pokyčiai veikia jų apsisprendimą rinkimuose;
- kaip rinkimų tipas (Seimo, Prezidento, Savivaldybių) keičia rinkėjų aktyvumą ir apsisprendimą;
- ar socialiniai ir demografiniai pokyčiai taip pat ir rinkimų tipas turi įtakos rinkėjų balsavimui?

Išsamūs Lietuvos rinkimų statistikos tyrimai turėjo atsakyti į klausimą, kodėl 1996 m. ir 2000 m. Seimo rinkimuose taip ryškiai pakito rinkėjų nuomonė (8.2 lentelė).

Prognozuojant rinkėjų elgesį ir jų pasirinkimą, taikomi daugialypės logistinės regresijos modeliai, kuriuose kaip nepriklausomi kintamieji naudojami šie socialiniai bei ekonominiai rodikliai:

- rasė,
- lytis,
- išsilavinimas,
- amžius,
- rinkimų tipas,
- pajamos,

- būsto nuomininkas/savininkas,
- laikas, kurį gyvena šiame būste,
- šeiminė padėtis,
- gyvenamosios vietovės padėtis.

8.2 lentelė. Partijos, politinės organizacijos ar koalicijos, surinkusios daugiausia balsų 1996 m. ir 2000 m. LR Seimo rinkimuose

Partija, politinė organizacija ar koalicija, 2000 m.	Balsų procentas	Partija, politinė organizacija ar koalicija, 1996 m.	Balsų procentas
A.Brazausko socialdemokratinė koalicija	31,08	Tėvynės sąjunga (Lietuvos konservatoriai)	29,80
Naujoji sąjunga (socialliberalai)	19,64	Lietuvos krikščionių demokratų partija	9,91
Lietuvos liberalų sąjunga	17,25	Lietuvos demokratinė darbo partija	9,52
Tėvynės sąjunga (Lietuvos konservatoriai)	8,62	Lietuvos centro sąjunga	8,24
Krikščionių demokratų sąjunga	4,19	Lietuvos socialdemokratų partija	6,60

Pavyzdys. Rinkėjų proporcijų lygybės tyrimai

Pirmiausia išnagrinėkime paprastą pavyzdį apie priešrinkiminiuosius viešosios nuomonės apklausos tyrimus. Rinkimų statistikoje dažnai tenka nagrinėti kintamuosius, išmatuotus nominalia skale. Pavyzdžiui, asmuo ketina dalyvauti rinkimuose arba neketina, ketina balsuoti už kažkurią partiją arba ketina balsuoti už kitą partiją ir t.t. Paprasčiausias atvejis, kai stebėjimai matuojami dichotomine skale, t. y., kai tyrimo rezultatas gali įgyti tik dvi reikšmes. Pavyzdžiui: išnagrinėkime Lietuvos Respublikos Seimo rinkimuose rinkėjų balsavimą už Lietuvos lenkų rinkimų akciją. Sutariame, kad reikšmė *vienas* bus požymis rinkėjų, kurie balsuoja už šią partiją ir *nulis* – kurie nebalsuoja. Gavome populiaciją vien tik iš nulių ir vienetų. Be abejonės, tai nėra normalusis skirstinys. Šiuo atveju stebimas binominis atsitiktinis dydis $B(I,p)$, kurio visos reikšmės sukongcentruotos į du taškus. Bet jei imties dydis n pakankamai didelis, nepaisant populiacijos formos, imties vidurkio skirstinys gali

būti aproksimuojamas normaliuoju skirstiniu (centrinė ribinė teorema). Galima manyti, kad imtis yra pakankamai didelė, jei $np > 5$ kai $p < q$; čia: p – balsavusių už Lenkų rinkimų akciją rinkėjų dalis, q – nebalsavusių rinkėjų dalis. Rinkimų statistikoje dažnai domimasi dviejų proporcijų skirtumu.

Tarkime, kad prieš 2000 m. Seimo rinkimus atliktos apklausos duomenimis, už Lietuvos lenkų rinkimų akciją iš 1345 apklaustų asmenų ketino balsuoti 28 asmenys. 2000 m. Seimo rinkimuose galiojančių biuletenių buvo 1 471 247, o už Lietuvos lenkų rinkimų akciją balsavo 28641 rinkėjai arba 1,95 proc. Norima nustatyti apklausos tikslumą, t. y., ar rinkėjų, ketinusių balsuoti už Lietuvos lenkų rinkimų akciją ir balsavusių proporcijos yra vienodos. Šį uždavinį galima spręsti keliais būdais. Išnagrinėkime tokio tipo uždavinių sprendimą, remdamiesi z kriterijumi.

1. Suformuluokime statistinę hipotezę:

$H_0: p_1 = p_2$, t.y proporcijos yra vienodos;

$H_1: p_1 \neq p_2$, t. y. proporcijos yra skirtingos.

2. Pasirenkame reikšmingumo lygmenį $\alpha = 0,05$.

3. Parenkame ir apskaičiuojame statistinį kriterijų. Hipotezei tikrinti naudosime z kriterijų, kadangi imtys yra didelės, ir binominį skirstinį galima aproksimuoti normaliuoju:

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)(1/n + 1/m)}}; \quad (8.1)$$

čia: $p_1 = k_1/n$; k_1 – ketinusių balsuoti rinkėjų skaičius, n – apklausoje dalyvavę rinkėjai; p_2 – 2000 m. už Lietuvos lenkų rinkimų akciją balsavusių rinkėjų skaičius, m – visi 2000 m. rinkimuose dalyvavę rinkėjai, $p = (k_1 + k_2)/(n + m)$.

Kai $p_1 = p_2$ šio kriterijaus reikšmės pasiskirsčiusios pagal normalųjį skirstinį, kurio vidurkis yra 0 ir standartinis nuokrypis yra 1. Normalusis skirstinys simetriškas nulinio atžvilgiu, todėl esant dvipusei alternatyvai $H_1: p_1 \neq p_2$ kritinės reikšmės yra $z_{\alpha/2}$. Čia $z_{\alpha/2}$ yra standartinio normaliojo skirstinio $\alpha/2$ lygmens kritinė reikšmė.

Apskaičiuojame:

$$p_1 = 28/1345 = 0,0208;$$

$$p_2 = 28641/1471247 = 0,0195;$$

$$p = (28 + 28641)/(1345 + 1471247) = 0,0195;$$

$$z = \frac{0,0208 - 0,0195}{\sqrt{0,0195(1 - 0,0195) \cdot 0,000744}} = \frac{0,0013}{0,00377} = 0,359.$$

4. Sprendimo priėmimas (išvadų formulavimas).

Kadangi alternatyva dvipusė, o parinktas reikšmingumo lygmuo $\alpha = 0,05$, norint priimti sprendimą, reikia žinoti $z_{0,025}$ reikšmę. Ją galima sužinoti iš statistinių lentelių arba iš statistinių duomenų analizės programų. *Skaičiuojant galima naudotis šiomis z_α reikšmėmis:*

$$z_{0,005} = 2,575; \quad z_{0,01} = 2,326 \quad z_{0,025} = 1,96; \quad z_{0,05} = 1,64; \quad z_{0,1} = 1,281.$$

Hipotezė H_0 atmetama, jei gauta reikšmė $|z| > z_{\alpha/2}$. Hipotezė H_0 neatmetama, jei gauta reikšmė $|z| \leq z_{\alpha/2}$. Vienu ar kitu būdu gauname, kad $z_{0,025} = 1,96$.

Kadangi apskaičiuota reikšmė $|z| = 0,359 \leq 1,96$, tai hipotezės H_0 atmesti negalima. Galima sakyti, kad apklausos rezultatai tokie pat, kaip rinkimų rezultatai.

8.3. Logistinė regresija

Tais atvejais, kai reikia nustatyti ryšį tarp vieno dichotominio kintamojo ir vieno ar kelių nepriklausomų kintamųjų, taikomi logistinės regresijos metodai. Logistinė regresija naudojama tada, kai priklausomas kintamasis gali įgauti tik dvi reikšmes. Pavyzdžiui, rinkėjas gali balsuoti už Naująją sąjungą, bet gali už ją ir nebalsuoti; studentas egzaminą gali išlaikyti, bet gali ir neišlaikyti; pacientas gali pasveikti arba nepasveikti; konkursą galima laimėti arba nelaimėti ir pan.

Kodėl norint šiuo atveju apskaičiuoti regresijos koeficientus, netinka standartiniai daugialypės regresijos metodai? Daugialypė regresija „nežino“, kad priklausomo kintamojo pobūdis yra dichotominis, todėl sudarytu modelių galima prognozuoti reikšmes, didesnes už 1 ir mažesnes už 0. Taigi standartiniai daugialypės regresijos metodai ignoruotų apribojimus (dichotomiškumą), keliamus priklausomam kintamajam.

Regresijos modelius galima sudaryti taip, kad, užuot prognozavus dvejetainį kintamąjį, bus prognozuojamas tolydusis kintamasis, kuris natūraliai kinta nuo 0 iki 1. Dažniausiai naudojami tokie modeliai yra du: logistinė regresija (*logit regression*) ir probistinė regresija (*probit regression*).

Logistinės regresijos modelis

Naudojant logistinę regresiją galima tirti ir prognozuoti priklausomo dichotominio kintamojo ir nepriklausomų kintamųjų, išmatuotų bet kuria skale (nepriklausomi kintamieji ir liekanos gali neturėti normaliojo skirstinio), ryšius. Prognozuojamos reikšmės niekada nebus didesnės (ir nebus lygios) už 1 ir nebus mažesnės (ir lygios) už 0. Tai pasiekama taikant logistinės regresijos modelį:

$$y = \frac{\exp(a + b_1x_1 + b_2x_2 + \dots + b_nx_n)}{1 + \exp(a + b_1x_1 + b_2x_2 + \dots + b_nx_n)} \quad (8.2)$$

Akivaizdu, kad ir kokio dydžio būtų x reikšmės, modelio prognozuojamos y reikšmės visada bus nuo 0 iki 1.

Pavadintas logistiniu (*logit*) šis modelis todėl, kad jis lengvai linearizuojamas taikant logistinę transformaciją. Kadangi priklausomas kintamasis y kinta nuo 0 iki 1, jis gali būti nagrinėjamas kaip tolydi tikimybė p , kuri kinta nuo 0 iki 1. Ši tikimybė gali būti transformuojama:

$$p' = \ln\left(\frac{p}{1-p}\right). \quad (8.3)$$

Ši transformacija vadinama logistine transformacija. Kadangi p' teoriškai gali įgyti bet kurią reikšmę nuo minus iki plus begalybės, tai šias transformuotas reikšmes jau galima naudoti įprastame tiesinės regresijos modelyje:

$$p' = a + b_1x_1 + b_2x_2 + \dots + b_nx_n + e. \quad (8.4)$$

8.4. Logistinė regresija *Statistica* programa

Išnagrinėkime fiktyvius duomenis (8.3 lentelė) apie rinkėjų amžių ir apsisprendimą balsuoti už ABC partiją (1 – apsisprendęs balsuoti, 0 – apsisprendęs nebalsuoti). Užduotis – įvertinti tikimybę, kad 50 metų ir vyresnio amžiaus rinkėjas balsuos už ABC partiją.

8.3 lentelė. Amžius ir apsisprendimas balsuoti

Amžius	60	40	52	56	38	65	50	44	34	53	32	52	26	42	30	19	36	38	42	52
Balsavimas	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0	0	0	0	0	0

STATISTICA programoje logistinės regresijos procedūra yra tokia:

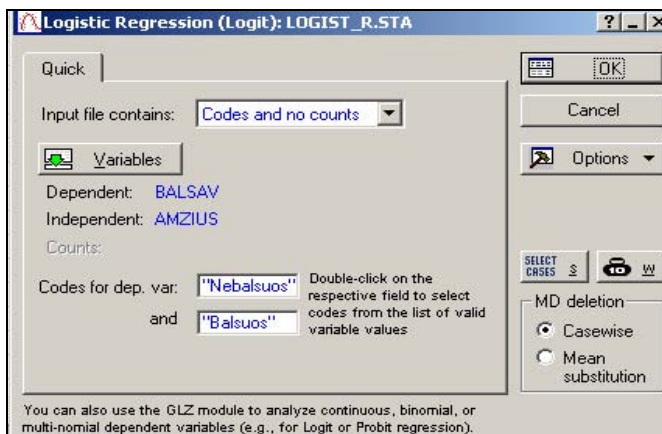
Statistics

Advanced Linear/Nonlinear Models

Nonlinear Estimation

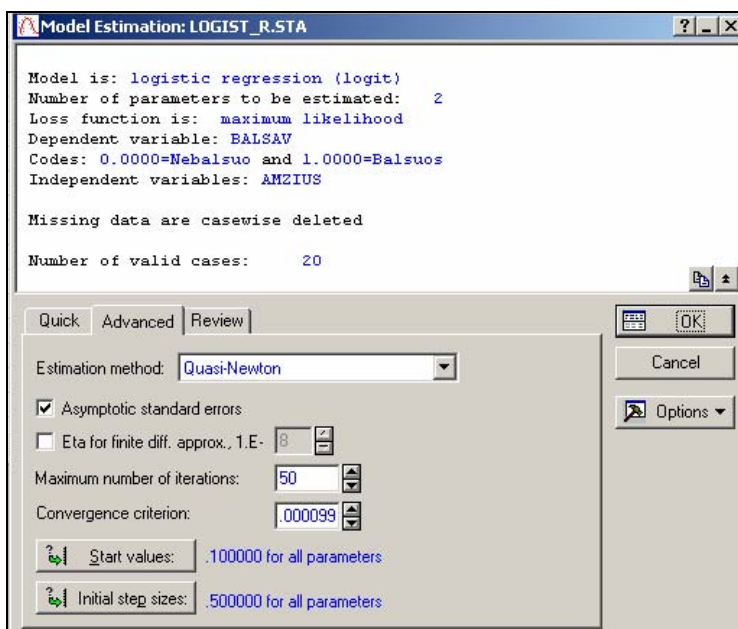
Quick Logit regression

Atskleistame duomenų įvedimo lange (8.1 pav.) nurodomas priklausomas dichotominis kintamasis (*BALSAV*) ir nepriklausomas kintamasis (*AMŽIUS*). Lange *Codes for dep.var* nurodomi priklausomo kintamojo kodai. Kodas, kurio reikšmė yra 0, rašomas pirmiausia.



8.1 pav. Logistinės regresijos duomenų įrašymo langas

Kai priklausomas kintamasis yra dichotominis, tai regresinės analizės prielaida, kad visų liekanų dispersija yra vienoda, netenkinama, todėl parametrams įvertinti taikomas ne mažiausių kvadratų, o maksimalaus tikėtino metodo. Paspaudus *OK* klavišą, *Advanced* lange galima pasirinkti iteracinį skaičiavimo metodą, konvergavimo kriterijų ir pradinės reikšmės (8.2 pav.).

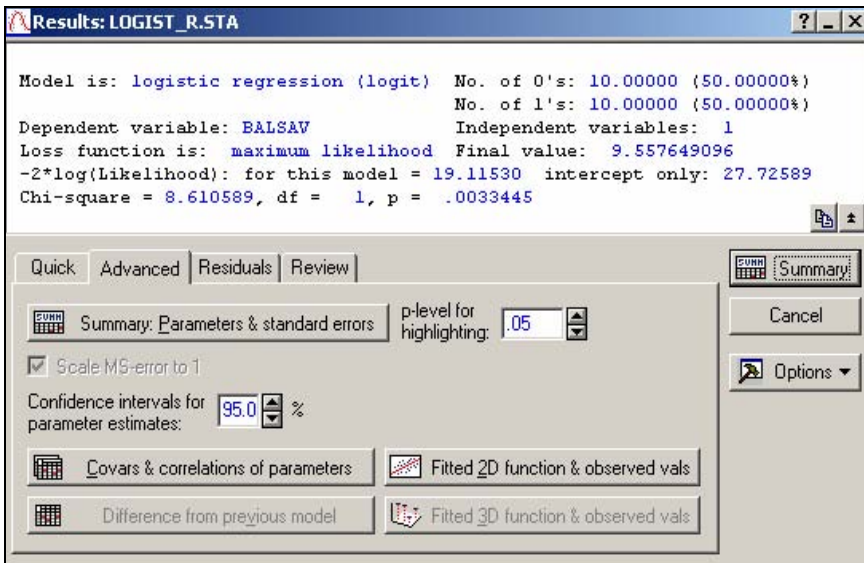


8.2 pav. Modelio parametų pasirinkimo dialoginis langas

Modelio parametų pasirinkimo lange galima nurodyti, kad įvertinus parametrus bus apskaičiuojami jų asimptotiniai standartiniai nuokrypiai (pasirinkimas *Asymptotic standard errors*). Pasirinkus modelio parametrus pradedamas iteracinis skaičiavimo procesas. Skaičiavimo rezultatai pateikiami rezultatų lange (8.3 pav.).

8.3 pav. pateikiama apskaičiuota χ^2 reikšmė, gauta lyginant pasirinktą modelį su modeliu, kurį sudaro tik regresijos lygties laisvasis narys (atkirtimo koeficientas *intercept*). Šiuo atveju χ^2 kriterijaus stebimasis reikšmingumo lygmuo yra $p=0,0033$; tai rodo, kad modelis yra statis-

tiškai reikšmingas. Taigi galima daryti išvadą, kad balsavimas susijęs su rinkėjo amžiumi.



8.3 pav. Logistinės regresijos rezultatų langas

Parametrų įverčiai gaunami paspaudus klavišą *Summary:Parameters & standard errors* (koeficientų paklaidos pateikiamos tik tuo atveju, kai pasirinkama *Asymptotic standard errors*; pažymėta 8.2 pav.).

8.4 lentelė. Logistinės regresijos koeficientų statistika

<i>Model: Logistic regression (logit) N of 0's:10 1's:10</i>		
<i>Dep. var: BALSAV Loss: Max likelihood (MS-err. scaled to 1)</i>		
<i>Final loss: 9,557649096 Chi²(1)=8,6106 p=0,00334</i>		
	Const.B0	AMŽIUS
<i>Estimate</i>	-6,568	0,152
<i>Standard Error</i>	2,965	0,068
<i>t(18)</i>	2,215	-2,253
<i>p-level</i>	0,040	0,037

8.4 lentelėje pateikti regresijos lygties koeficientai, jų asimptotiniai standartiniai nuokrypiai, kriterijaus t reikšmės, kurios gaunamos padalijus koeficientų įverčius iš jų standartinių nuokrypių, ir kiekvieno koeficiento statistinis reikšmingumas (p -level). Abu koeficientai yra statistiškai reikšmingi, kai reikšmingumo lygmuo $\alpha=0,05$.

Rezultatų interpretavimas

Gauti koeficientai interpretuojami taip pat, kaip ir tiesinės regresijos. Galima parašyti regresijos lygtį:

$$\hat{p}' = -6,568 + 0,152 \cdot \text{amžius} .$$

Šios lygties parametrai susiję su $\ln(p/(1-p))$, o ne su balsavimo tikimybėmis. Norint apskaičiuoti balsavimo tikimybę konkrečiu atveju, pritaikoma (8.2) formulė. Pavyzdžiui, kai amžius=50 metų, balsavimo tikimybė apskaičiuojama taip:

$$\begin{aligned} \hat{p}' &= -6,568 + 0,152 \cdot 50 = -1,032; \\ \hat{p} (50) &= \exp(1,032) / (1 + \exp(1,032)) = (2,81/3,81) = 0,737. \end{aligned}$$

Prognozavimas ir klasifikavimas

Klavišu *Observed, predicted, residual vals* prognozuojamos reikšmės pateikiamos kaip balsavimo tikimybės (žr. 8.5 lentelę).

8.5 lentelė. Prognozuojama balsavimo už ABC partiją tikimybė

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
0,93	0,38	0,79	0,88	0,31	0,97	0,74	0,53	0,2	0,82	0,16	0,79	0,07	0,46	0,12	0,02	0,25	0,31	0,46	0,79

Norint gautas reikšmes naudoti klasifikuojant, reikia sudaryti klasifikavimo taisyklę.

Klasifikavimo taisyklė:

jeigu gauta tikimybės reikšmė didesnė už 0,5, tai prognozuojama, kad y reikšmė yra 1;

jeigu gauta tikimybės reikšmė didesnė už 0,5, tai prognozuojama, kad y reikšmė yra 0; tikimybės reikšmė 0,5 galima tiek vienu, tiek kitu atveju.

Duomenų klasifikavimas. Klavišu *Classification of cases & odds ratio* sudaroma šio modelio teisingai ir neteisingai klasifikuotų atvejų lentelė (8.6 lentelė).

8.6 lentelė. Balsavimo atpažinimo klasifikacinė lentelė

<i>Classification of Cases</i>			
<i>Odds ratio: 9,333</i>			
	<i>Pred.</i>	<i>Pred.</i>	<i>Percent</i>
	Nebalsuos	Balsuos	<i>Correct</i>
Nebalsuos	8	2	80
Balsuos	3	7	70

Apskaičiuotos tikimybės, didesnės nei 0,5, klasifikuojamos kaip balsavimas „už“, o mažesnės nei 0,5 – balsavimas „prieš“. Jei teisingai ir neteisingai klasifikuotų atvejų santykis (*Odds ratio*) didesnis už 1, vadinasi, klasifikacija yra geresnė nei grynai atsitiktinis spėjimas. Tačiau svarbu atsiminti, kad parametrai buvo apskaičiuoti taip, kad maksimizuotų stebimų duomenų tikimybę. Taigi iš tikrųjų modelio tinkamumą geriausia iširti pritaikius jį naujiems stebėjimams.

Užduotys

8.1. Atliekant tyrimą apie X partijos reklamos įtaką rinkėjų apsisprendimui, 25 asmenys nurodė vidutinį laiką per savaitę, kai jie žiūri TV rinkimams skirtas laidas. Po rinkimų tų pačių asmenų buvo klausama, ar jie balsavo už X partijos kandidatą. Tyrimo rezultatai pateikti 8.7 lentelėje. Nustatykite, ar veiksminga buvo partijos reklama, ar tiems rinkėjams, kurie ilgiau reklamą stebėjo, ji padėjo apsispręsti už X partijos kandidatą balsuoti.

8.7 lentelė. Reklamos įtaka rinkėjų apsisprendimui.

Žiūrėta rinkimų laidų (val.)	4	8	2	6	5	1	4	3	5	2	8	3	8
Balsavo už partijos kandidata	Ne	Ne	Ne	Taip	Taip	Ne	Ne	Ne	Taip	Ne	Taip	Ne	Taip
Žiūrėta rinkimų laidų (val.)	2	5	3	2	8	6	4	5	1	7	6	2	
Balsavo už partijos kandidata	Ne	Taip	Ne	Ne	Taip	Ne	Taip	Ne	Ne	Taip	Taip	Taip	

8.2. 8.8 lentelėje pateikti duomenys apie gyventojų lytį, jų gyvenimo kokybę (0 – labai bloga, 100 – labai gera) ir nusiteikimą dalyvauti rinkimuose. Įvertinkite tikimybę, kad moteris, gyvenimo kokybę vertinanti 90 balų, nedalyvaus rinkimuose.

8.8 lentelė. Duomenys apie gyventojų lytį, gyvenimo kokybę ir nusiteikimą dalyvauti rinkimuose

Lytis	mot.	mot.	mot.	vyr.	vyr.	vyr.	mot.	mot.	mot.	vyr.
Ar dalyvaus rinkimuose	taip	taip	taip	taip	taip	taip	taip	taip	taip	taip
Gyvenimo kokybė	68	52	56	76	96	80	56	48	44	80
Lytis	vyr.	vyr.	vyr.	mot.	vyr.	mot.	mot.	mot.	mot.	mot.
Ar dalyvaus rinkimuose	ne	ne	ne	ne	ne	ne	ne	ne	ne	ne
Gyvenimo kokybė	40	56	16	16	60	56	16	32	20	28

Pastaba. Kai logistinės regresijos modelio pagalba nagrinėjamas daugiau nei vienas nepriklausomas kintamasis, pirmiausia galima įvertinti modelį su vienu kintamuoju, o po to – jau su dviem kintamaisiais. Norint sėkmingai atlikti analizę, gali tekti papildomai įtraukti ar pašalinti vieną ar daugiau kintamųjų.



LITERATŪRA

1. Aksomaitis A. Tikimybių teorija ir statistika. Kaunas: Technologija, 2000.
2. Atkinson A.B. On the measurement of inequality. *Journal of Economic Theory*. 1970, 2, 244–263 p.
3. Alonso W., Starr P. *The Politics of Numbers*. New York: Russel Sage Foundation, 1987.
4. Bauman Z. *Globalizacija: pasekmės žmogui*. Vilnius: Strofa, 2002.
5. Chadjipadelis T. Using Statistics as a tool in political research. The case of electoral behavior. *ICOTS6*, p.1–6, 2002.
6. Chiang, Chin Long. *The Life Table and Its Applications*. Robert E. Krieger Publishing Company, Inc. Malabar, Fl, 1984.
7. Cowell FA. *Measuring Inequality*. Oxford: Philip Allan, 1977.
8. Социальная статистика. Под редакцией И.И.Елисеевой. Москва: Финансы и статистика, 2003.
9. Čekanavičius V., Murauskas G. *Statistika ir jos taikymai I*. Vilnius: TEV, 2000.
10. Duncan J.W, Gross A.C. *Statistics for the 21st century*. Irwin:The Dun and Breadstreet Co, 1995.
11. Edwards D. and Hamson M. *Guide to Mathematical Modelling*. Boca Raton, Fla.:CRC Press, 1989.
12. Ефимова М.Р., Петрова Е.В., Румянцев В.Н. *Общая теория статистики*. Москва: ИНФРА–М, 1999.
13. Эйдintas А. *Литовская эмиграция в страны Северной и Южной Америки в 1868–1940 г.* Вильнюс: Мокслас, 1989.
14. Francis A. *Business Mathematics and Statistics*. London:DP Publications, 1993.
15. Giere, Ronald N. *Science Without Laws*. Chicago: University of Chicago Press, 1999.
16. Goffman E. *Strategic Interaction*. Oxford: Basil Blackwell, 1970.
17. Gudavičiūtė-Bartosevičienė V. *Ekonominė statistika II d*. Kaunas: Technologija, 1997.
18. Hernes, Gudmund. The process of entry into first marriage. *American Sociological review* 1972, 37, 173–182 p.

19. Kaplan GA, Pamuk ER, Lynch JW, Cohen RD, Balfour JL. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *British Medical Journal*, 1966, 312, 999–1003 p.
20. Kawachi I, Kennedy BP. The relationship of income inequality to mortality – Does the choice of indicator matter? *Soc Sci Med* 1997, 45, 1121–1127 p.
21. Kennedy BP, Kawachi I, Glagg R, Prothrow-Stith. Income distribution, socioeconomic status, and self-rated health: A US multi-level analysis. *Br Med J* 1998, 317, 917–921 p.
22. Klosterman, R.E. *Community and Analysis Planning Techniques*. Rowmand and Littlefield Publishers, Inc. Savage, Maryland, 1990.
23. Lietuvos statistikos metraštis. Vilnius: Statistikos departamentas, 2001.
24. Лунеев В.В. Юридическая статистика. Москва: Юрист, 1999.
25. Martišius S. ir kt. *Ūkio statistika (teorijos ir praktikos apybraižos)*. Vilnius: LB, 1995.
26. Morgenstern, O. *On the Accuracy of Economic Observations*. 2nd ed. Princeton, NJ: Princeton University Press, 1963.
27. Patton, M.Q. *Qualitative evaluation and research methods*. Newbury Park, CA: SAGE Publications, 1990.
28. Rudzkiene V. *Statistinės technologijos teisėje ir valdyme*. Vilnius:LTU, 2003.
29. Sakalauskas V. *Statistika su Statistica*. Vilnius: Margi raštai, 1998.
30. Sen A. *On Economic Inequality*. Oxford: Oxford University Press, 1973.
31. Schwarze J. *Grundlagen der Statistik 2. 5th edition*; Herne/Berlin: Neue Wirtschaftsbriefe, 1993.
32. Soobader M.J., LeClere F.B. Aggregation and the measurement of income inequality: effects on morbidity. *Social Science & Medicine* 1999, 48, 733–744 p.
33. Vaitekūnas S. *Lietuvos teritorija ir gyventojai*. Vilnius: Alma Littera, 1998.
34. Wilkinson RG. *Unhealthy Societies. The Afflictions of Inequality*. London:Routledge, 1996.
35. Visuotinė žmogaus teisių deklaracija. Teisės problemos. Vilnius: Teisės institutas, 1999. Nr.1–2.

36. Voter Participation in Canada: Is Canadian Democracy in Crisis? Center for Research and Information on Canada, October, 2001.
37. Ziliak S.T. Some Tendencies of Social Welfare and the Problem of Interpretation. *Cato Journal*, Vol.21, No3, 499–513 p.
38. Statistikos departamentas prie Lietuvos Respublikos Vyriausybės: <http://www.std.lt/>
39. Nusikalstamumo prevencijos Lietuvoje centras: <http://www.nplc.lt/>
40. Statistical Resources on the Web Political Science: <http://www.lib.umich.edu/govdocs/stpolisc.html>
41. The World Factbook: <http://www.odci.gov/cia/publications/factbook/index.html>
42. Lietuvos rinkimų statistika: <http://www3.lrs.lt/n/rinkimai/20001008/rdl.htm-13.htm>
43. Advanced Sample Size Calculator: <http://www.questaresearch.com/>



PRIEDAI

A lentelė. Stjudento skirstinio α lygmens kritinės reikšmės $t_{\alpha}(n)$

n/ α	$\alpha=0,40$	$\alpha=0,10$	$\alpha=0,05$	$\alpha=0,025$	$\alpha=0,01$	$\alpha=0,005$
1	0,325	3,078	6,314	12,706	31,821	63,657
2	0,289	1,886	2,920	4,303	6,965	9,925
3	0,277	1,638	2,353	3,182	4,541	5,841
4	0,271	1,533	2,132	2,776	3,747	4,604
5	0,267	1,476	2,015	2,571	3,365	4,032
6	0,265	1,440	1,943	2,447	3,143	3,707
7	0,263	1,415	1,895	2,365	2,998	3,499
8	0,262	1,397	1,860	2,306	2,896	3,355
9	0,261	1,383	1,833	2,262	2,821	3,250
10	0,260	1,372	1,812	2,228	2,764	3,169
11	0,260	1,363	1,796	2,201	2,718	3,106
12	0,259	1,356	1,782	2,179	2,681	3,055
13	0,259	1,350	1,771	2,160	2,650	3,012
14	0,258	1,345	1,761	2,145	2,624	2,977
15	0,258	1,341	1,753	2,131	2,602	2,947
16	0,258	1,337	1,746	2,120	2,583	2,921
17	0,257	1,333	1,740	2,110	2,567	2,898
18	0,257	1,330	1,734	2,101	2,552	2,878
19	0,257	1,328	1,729	2,093	2,539	2,861
20	0,257	1,325	1,725	2,086	2,528	2,845
21	0,257	1,323	1,721	2,080	2,518	2,831
22	0,256	1,321	1,717	2,074	2,508	2,819
23	0,256	1,319	1,714	2,069	2,500	2,807
24	0,256	1,318	1,711	2,064	2,492	2,797
25	0,256	1,316	1,708	2,060	2,485	2,787
30	0,256	1,310	1,697	2,042	2,457	2,750
40	0,255	1,303	1,684	2,021	2,423	2,704
60	0,254	1,296	1,671	2,000	2,390	2,660
120	0,254	1,289	1,658	1,980	2,358	2,617

B lentelė. χ^2 skirstinio α lygmens kritinės reikšmės

n/ α	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,025$	$\alpha = 0,01$
1	2,71	3,84	5,02	6,63
2	4,61	5,99	7,38	9,21
3	6,25	7,81	9,35	11,34
5	9,24	11,07	12,83	15,09
6	10,64	12,59	14,45	16,81
7	12,02	14,07	16,01	18,48
8	13,36	15,51	17,53	20,09
9	14,68	16,92	19,02	21,67
10	15,99	18,31	20,48	23,21
12	18,55	21,03	23,34	26,22
13	19,81	22,36	24,74	27,69
14	21,06	23,68	26,12	29,14
15	22,31	25,00	27,49	30,58
16	23,54	26,30	28,85	32,00
18	25,99	28,87	31,53	34,81
19	27,20	30,14	32,85	36,19
20	28,41	31,41	34,17	37,57
21	29,62	32,67	35,48	38,93
22	30,81	33,92	36,78	40,29
23	32,01	35,17	38,08	41,64
24	33,20	36,42	39,36	42,98
25	34,38	37,65	40,65	44,31
26	35,56	38,89	41,92	45,64
27	36,74	40,11	43,19	46,96
28	37,92	41,34	44,46	48,28
29	39,09	42,56	45,72	49,59
30	40,26	43,77	46,98	50,89
40	51,80	55,76	59,34	63,69
50	63,17	67,50	71,42	76,15
60	74,40	79,08	83,30	88,38
70	85,53	90,53	95,02	100,42
80	96,58	101,88	106,63	112,33
90	107,56	113,14	118,14	124,12
100	118,50	124,34	129,56	135,81

Doc. dr. Rudzkienė, Vitalija

Ru44 Socialinė statistika: vadovėlis. – Vilnius: Mykolo Romerio universiteto Leidybos centras, 2005. – 260 p., 15 paveikslų, 100 lentelių, 42 grafikai.

Bibliogr.: p. 253–255.

ISBN 9955-19-002-7

Vadovėlyje nagrinėjamos svarbiausios socialinės statistikos sritys, šių sričių būdingos problemos ir tyrimų metodologija. Pateikiami tipinių uždavinių sprendimo būdai, statistinių išvadų pagrįstumo analizės ir rezultatų interpretavimo pavyzdžiai. Realiais duomenimis pagrįsta demografinės statistikos, gyventojų gerovės ir pajamų diferenciacijos, nedarbo, integruotų rodiklių, teisinės ir politinės statistikos sričių analizė padeda suvokti statistikos metodų taikymo galimybes vertinant socialinius procesus ir numatant jų kitimo tendencijas.

Vadovėlis skirtas viešojo administravimo, teisės ir valdymo bei kitų socialinių mokslų studentams. Jo tikslas – ugdyti profesionalų analitinį mąstymą, formuoti gebėjimus analizuoti ir prognozuoti socialinius reiškinius.

UDK 311:316(075.8)

Vitalija Rudzkienė
SOCIALINĖ STATISTIKA
Vadovėlis

Redaktorė *Genovaitė Gudonienė*
Maketuotoja *Regina Bernadišienė*
Viršelio autorė *Stanislava Narkevičiūtė*

SL 585. 2005 04 05. 12,56 leidyb. apsk. l.

Tiražas 1000 egz. Užsakymas .

Išleido Mykolo Romerio universiteto Leidybos centras, Ateities g. 20, LT-08303 Vilnius.

Tinklapis internete www.mruni.lt

Elektroninis paštas leidyba@mrni.lt

Spausdino UAB „Baltijos kopija“, Kareivių g. 13 b, Vilnius.

Tinklapis internete www.kopija.lt

El. paštas info@kopija.lt