

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro darbas

**VERSLININKŲ PASITENKINIMĄ DARBU ĮTAKOJANČIŲ
VEIKSNIŲ DAUGIAMATĖ ANALIZĖ**

Čepulinskaitė Laura

VILNIUS 2007

Ekonometrinės analizės katedra

Darbo vadovas **doc. dr. M. Radavičius** _____
(parašas)

Darbas apgintas 2007 m. birželio mėn. 4 d.

Registravimo Nr. _____
2007-05-24 _____

TURINYS

SANTRAUKA	5
SUMMARY	6
ĮVADAS.....	7
I. TEORINIS-METODINIS SKYRIUS.....	9
I.1.Pagrindinės sąvokos	9
I.1.1 Kategoriniai kintamieji.....	9
I.1.2. Optimalusis skalės transformavimas.....	9
I.1.3. Multikolinearumas.....	10
I.1.4. Polichorinės ir tetrachorinės koreliacijos.....	10
I.2. Faktorinė analizė.....	12
I.2.1. Matematinis modelis.....	12
I.2.2. Duomenų tikimo faktorinei analizei matai.....	13
I.2.3. Faktorių išskyrimas pagrindinių komponentų analizės metodu.....	13
I.3. Apibendrintasis Tiesinis modelis.....	15
I.3.1. Apibendrintojo Tiesinio modelio elementai	15
I.3.2 Logistinė regresija.....	17
I.3.2.1. Statistinės išvados.....	17
I.3.3. Daugiareikšmis (<i>multinomial</i>) logit modelis.....	18
I.3.4. Ranginė (<i>ordinal</i>) regresija.....	19
I.3.4.1. Lygiagrečiųjų tiesių (plokštumų) prielaida.....	19
I.3.4.2. Proporcionaliųjų galimybių modelis.....	20
I.3.4.3. Proporcionaliųjų rizikų modelis.....	21
II. TIRIAMASIS- ANALITINIS SKYRIUS.....	22
II.1. Duomenys.....	22
II.1.1. Duomenų šaltinis ir apklausos dalys.....	22
II.1.2. Kintamieji ir jų transformacijos.....	22
II.1.3. Trumpa imties apžvalga.....	22
II.1.4. Multikolinearumo tyrimas.....	23
II.2. Faktorinė analizė.....	25
II.3. Modelio parinkimas.....	26

IŠVADOS	28
LITERATŪRA	29
PRIEDAI.....	30

SANTRAUKA

Šis darbas skirtas Italijos verslininkų pasitenkinimo darbu analizei, naudojantis kokybinių požymių daugiamačių statistikos metodais.

Duomenys yra iš Italijos Veneto regiono verslininkų tyrimo, kurį 2006 metais pradėjo Padujos universiteto Statistikos departamentas. Imtį sudaro 1216 stebėjimų iš beveik 113000 Veneto regiono (Šiaurės rytų Italija) verslininkų populiacijos.

Priklausomasis kintamasis, t.y. pasitenkinimas darbu, išmatuotas ranginėje keturių taškų Likert skalėje su šiomis kategorijomis: „nepatenkintas“, „nei patenkintas, nei nepatenkintas“, „pakankamai patenkintas“ ir „patenkintas“. Paaiškinamuosius kintamuosius sudaro demografinės, firmos charakteristikos, tapimo verslininku motyvus, darbą, laisvalaikį ir ateities perspektyvas atspindintys kintamieji.

Beveik visi kintamieji, taip pat ir priklausomasis, yra kategoriniai. Todėl pagrindinis šio darbo tikslas yra išsirinkti geriausiai tokio tipo duomenims tinkantį modelį. Aptariama galimybė naudotis tiesine regresija, atlikus optimalųjį matavimo skalės transformavimą (optimal scaling), logistine regresija, prieš tai apjungus priklausomojo kintamojo rangus, daugianariu logit modeliu, analizuojant kiekvienos kategorijos tikimybę atskirai ir rangine regresija su logit ir cloglog sąryšio funkcijomis.

Potencialių pasitenkinimo darbu veiksmų skaičius- pakankamai didelis (daugiau negu 30), todėl itin svarbus uždavinys yra parinkti aiškinančiuosius kintamuosius, vengiant multikolinearumo problemos. Siekiant sumažinti kintamųjų skaičių apsvarstyta galimybė pritaikyti faktorinę analizę. Kadangi ji, kaip taisyklė, taikoma tik kiekybiniam duomenims, pasiūlytas alternatyvus faktorių išskyrimo būdas vietoj Pearson'o koreliacijų įstatant polichorinių bei tetrachorinių koreliacijų koeficientus ranginiams ir dvireikšmiams kintamiesiems.

Rezultatai rodo, kad atsižvelgiant į modelio prielaidų tenkinimą, modelio tikimo duomenims statistikas ir klasifikacijos lenteles duomenims geriausiai tinka ranginė regresija su cloglog sąryšio funkcija. Multikolinearumo analizės dėka nustatyti stipriai susiję kintamieji buvo transformuoti arba neįtraukti į modelius. Pritaikius faktorinę analizę kintamųjų skaičių pavyko sumažinti iki trijų faktorių, kuriuos, atsižvelgiant į su jais koreliuotus kintamuosius, galima sąlyginai pavadinti ‚darbo‘, ‚laisvalaikio‘ ir ‚pažangios firmos‘ identifikatoriais. Šie faktoriai buvo panaudoti tolimesnėje analizėje kaip aiškinamieji kintamieji vietoje su jais susijusių kintamųjų.

Pasitenkinimo darbu veiksniai įvertinti atskirai vyrams ir moterims. Iš rezultatų galima daryti išvadą, jog tiek ‚laisvalaikio‘, tiek ‚pažangios firmos‘ faktoriai teigiamai veikia abiejų lyčių atstovų pasitenkinimą. Tuo tarpu ‚darbo laiko‘ faktorius didina moterų ir mažina vyrų pasitenkinimą darbu. Daugiau patenkinti darbu yra nevedę arba išsiskykę, vyresni ir turintys aukštąjį išsilavinimą vyrai bei ištekėjusios, jaunesnės, neturinčios mokyklinio amžiaus vaikų, aukštesnįjį išsilavinimą įgijusios moterys. Verta paminėti, kad analizuojant priežastis, paskatinusias pradėti verslą, lyginant su bazine ‚savarankiško gyvenimo pradėjimo‘ kaip tapimo verslininku priežastimi, profesinė realizacija yra teigiamai susijusi su pasitenkinimu darbu abejoms lytims. Tuo tarpu kai verslo pradėjimo priežastis yra pinigai, moterims pasitenkinimas darbu auga, o vyrams- mažėja.

Raktiniai žodžiai: pasitenkinimas darbu, kategoriniai duomenys, ranginis priklausomasis kintamasis, logistinė regresija, multinomial logit, optimal scaling, ranginė regresija, logit, cloglog, multikolinearumas, polichorinės ir tetrachorinės koreliacijos, faktorinė analizė, lytis, Italija.

SUMMARY

The paper presents job satisfaction analysis among Italian entrepreneurs using multivariate statistical techniques.

Data are taken from Veneto region (North-East part of Italy) businessmen research started by Department of Statistics of Padova University at 2006. Sample consists of 1216 observations (real population is almost 113 000 entrepreneurs at this region).

The outcome variable for the job satisfaction is measured on an ordered, categorical, four-point Likert scale – „dissatisfied“, „neither satisfied nor dissatisfied“, „quite satisfied“ and „satisfied“. Explanatory variables include demographic items, firm characteristics, variables representing the reasons of having started the own business, items associated with work, leisure activities and future perspectives.

Quite all variables, as well as the dependent one, being categorical, the main objective of this work is to select an appropriate model for such type of data among possible alternatives. Discussions are made on possibility to apply linear regression with optimal scaling, binary logit for dichotomized dependent variable, multinomial logit for analysis of every single category and ordinal logistic regression with several link functions (logit and cloglog).

The number of possible determinants of job satisfaction being quite large (there were more than 30 questionnaire items associated with job satisfaction) it was of a great importance choosing which explanatory variables should be included in the models, considering the multicollinearity challenge. In order to reduce the number of variables, it was considered possibility to apply factor analysis. The latter being applied only for a quantitative data, an alternative method to extract factors was proposed evaluating polychoric and tetrachoric correlations for ordinal and dichotomous variables instead of Pearson correlations .

The results show that considering model assumptions, fit statistics and classification results, the most appropriate model for our data is ordinal logistic regression with cloglog link. Multicollinearity analysis was of a great help to detect some highly associated variables that were transformed or omitted of models. Applying factor analysis the quantity of variables was reduced detecting three factors, that according to the correlated with them variables were named as ‚work‘ ‚leisure‘ and ‚advanced firm‘ factors. Later they all were included in considered models as explanatory variables instead of the variables they represent.

Job satisfaction determinants were evaluated by gender. Results show that factors representing leisure and advanced firm both increase job satisfaction for men and women, while working time positively affects women job satisfaction and has negative influence for men job satisfaction. More satisfied with job seem to be unmarried or divorced, high-educated, older men and younger, married women who have not school age children and who have graduated superior school. It is worth mentioning that analysing reasons of becoming entrepreneur it is seen that ‚professional realization‘ affects positively job satisfaction for both genders comparing with the reference- ‚starting the autonomous life‘. While ‚money‘ as a reason of haven started the business has positive influence for women and negative one for men comparing with ‚starting autonomous life‘.

Key words: job satisfaction, categorical data, ordinal dependent variable, logistic regression, multinomial logit, optimal scaling, ordinal regression, logit, cloglog, multicollinearity, polychoric and tetrachoric correlations, factor analysis, gender, Italy.

IVADAS

Pasitenkinimas darbu – tai individo nuostata į savo darbą arba dar kitaip - subjektyvus savo karjeros suvokimas. Pati sąvoka "pasitenkinimas darbu" pradėta vartoti neseniai, XX a. antrojoje pusėje. Tuomet imtos kurti ir įvairios pasitenkinimo darbu teorijos, pradėti analizuoti su pasitenkinimu darbu susiję veiksniai.

Tiksliai apibrėžti, nuo ko priklauso pasitenkinimas darbu yra gana sudėtinga. Dažniausiai veiksniai, darantys įtaką asmens pasitenkinimui darbu skirstomi į išorinius ir vidinius. Vidiniai pasitenkinimą darbu įtakoiantys veiksniai yra susiję su asmeninėmis savybėmis, tokiais kaip lytis, išsilavinimas, amžius, šeimyninė padėtis. Tuo tarpu išoriniai veiksniai- tai firmos savybės, darbo ypatybės, uždarbis, streso faktoriai. Moterims pasitenkinimą darbu dažnai įtakoja ir gebėjimas suderinti šeimyninį gyvenimą ir darbo pareigas.

Paprastai yra atliekami darbuotojų pasitenkinimo darbu tyrimai, užsakyti įmonių, kurios, siekdamos išsilaikyti konkurencinėje kovoje, siekia nustatyti pagrindinius pasitenkinimo darbu veiksnius ir jį kelti tarp savo darbuotojų.

Ne mažiau įdomesnė, nors praktiškai mažiau naudinga ir todėl retai taikoma yra savo verslą sukūrusių žmonių pasitenkinimo darbu analizė. Kas gi galėtų lemti individo, esančio ant aukščiausio firmoje karjeros laiptelio ir galinčio kontroliuoti nemažą dalį ankščiau minėtų išorinių veiksnių, pasitenkinimą darbu? Sunku atsakyti į šį klausimą, nes visos pasitenkinimo darbu teorijos yra skirtos firmos darbuotojams, o ne jos savininkams ir bent jau kol kas nėra apibrėžti veiksniai, nuo kurių galėtų priklausyti pačių verslininkų pasitenkinimas darbu. Todėl mano darbo tikslas- bent iš dalies užpildyti šią spragą ir pabandyti atsakyti į ankščiau kilusį klausimą apie galimus pasitenkinimo darbu veiksnius, o taip pat ištirti, ar skiriasi verslininkų vyrų ir moterų pasitenkinimas darbu ir jį įtakoiantys faktoriai.

Verslininkų pasitenkinimui ir jo veiksniams analizuoti panaudoti duomenys yra iš Italijos Padujos universiteto Statistikos departamento 2006 m. pradėto Veneto regiono verslininkų tyrimo. Jo metu buvo apklausta 1216 verslininkų, iš kurių 826 moterys ir 390 vyrų. Šiame regione bendra verslininkų populiacija yra apie 113 000 verslininkų, kurių didžioji dauguma- smulkaus verslo atstovai. Tad ir išvadas apie pasitenkinimą darbu reikėtų daryti smulkiesiems verslininkams, nes galimas dalykas, kad didelių firmų savininkams pasitenkinimą darbu lemia kiek kitokie veiksniai ar bent jau skiriasi tų pačių veiksnių įtaka.

Pasitenkinimas darbu išmatuotas 4 taškų *Likert skalėje*. Tai viena populiariausių psichometrinių atsakymų skalė, kur atsakantysis apibrėžia savo sutikimo su teiginiu lygį. Paprastai Likert skalės duomenys traktuojami kaip ranginiai, ypač kai kategorijų yra nedaug, nes atsakantysis gretimus taškus suvokia kaip vienodai nutolusius. Mūsų atveju Likert skalę sudaro šios kategorijos: „nepatenkintas“, „nei patenkintas, nei nepatenkintas“, „pakankamai patenkintas“ ir „patenkintas“.

Kaip galimi pasitenkinimo darbu veiksniai analizuotos demografinės apklausiamojo charakteristikos, jo turimos firmos ypatybės bei veiklos rodikliai, verslo pradėjimo motyvai ir aplinkybės, darbo laikas, laisvalaikis ir ateities perspektyvos. Didžioji dauguma kintamųjų yra ranginiai arba dvireikšmiai, o atsižvelgus į prasmingumą ir paprastesnį interpretavimą, nuspręsta suskirstyti į kategorijas ir kai kuriuos kiekybinius kintamuosius, tokius kaip metai, darbuotojų skaičius, darbo ir poilsio valandų per dieną

bei atostogų per metus skaičius. Tad galiausiai visi analizei skirti kintamieji virto kokybiniais (arba kategoriniais).

Kategorinių duomenų analizei dažniausiai naudojamos dvimatės procedūros- kontingencijos lentelių analizė, asociacijos koeficientai. Daugiamatiai metodai pradėti naudoti visai neseniai, paskutiniaisiais dešimtmečiais smarkiai išaugus instrumentų kokybiniais duomenims analizuoti poreikiui. Tad tai- nauja, bet smarkiai vystoma statistikos ir ekonometrijos sritis.

Vienas iš pagrindinių mano darbo uždavinių- atsižvelgiant į duomenų specifiką parinkti geriausiai juos modeliuojantį metodą. Kai priklausomasis kintamasis yra ranginis, siūlomos šios įprastinių modelių alternatyvos: a) pritaikius optimalaus ranginės skalės transformavimo į kiekybinę skalę techniką (optimal scaling) naudoti paprastą tiesinę regresiją; b) apjungti rangus į dvi reikšmes ir naudoti logistinę regresiją; c) modeliuoti kiekvienos kategorijos (rango) tikimybę atskirai naudojant daugianarį (multinomial) logit modelį; d) patikrinus lygiagrečiųjų tiesių (parallel lines) prielaidą naudoti ranginę (ordinal) regresiją pasirenkant duomenims tinkančią sąryšio funkciją. Darbe bus aptarti visi šie modeliai, jų prielaidos, palygintos tinkamumo duomenims statistikos ir pagrįstos atitinkamo modelio pasirinkimo priežastys.

Labai dažnai, turint ranginius ar dvireikšmius duomenis, kyla multikolinearumo problemų. *Multikolinearumas* suprantamas kaip stiprios priklausomybės egzistavimas tarp dviejų (tokiu atveju vadinama tiesiog kolinearumu) ar daugiau paaiškinamųjų kintamųjų. Multikolinearumo pagrindinė pasekmė – regresijos funkcijos koeficientų ‚nepatikimumas‘ (tiek dydžio, tiek ženklų prasme), aukštos standartinės klaidos, kurios net ir svarbius analizei kintamuosius paverčia statistiškai nereikšmingais. Todėl vienas iš mano uždavinių sudarant modelį- tinkamai parinkti paaiškinamuosius kintamuosius, vengiant multikolinearumo problemos.

Reikšmingų paaiškinamųjų kintamųjų gausa modelyje paprastai yra teigiamas dalykas, padedantis geriau suvokti tiriamojo objekto veiksmų visumą. Tačiau pasitaiko, kad grupė kintamųjų atspindi kažkokį bendresnį latentinį veiksnį, todėl interpretacijos palengvinimo ir modelio supaprastinamo dėlei susijusius kintamuosius galima apjungti išskiriant pagrindinius faktorius (būtent juos galima interpretuoti kaip latentinius kintamuosius). Tokio uždavinio poreikis kilo ir atliekant pasitenkinimo darbu veiksmų analizę, nes reikšmingų kintamųjų skaičius siekė beveik 30 ir dauguma jų pagal reikšmę ir įtaką buvo tarsi pasiskirstę grupelėmis.

Problema- faktorinė analizė paprastai taikoma tik kiekybiniais kintamiesiems, nes faktoriai išskiriami pasinaudojant Pearson‘o koreliacijų matrica. Ranginių ir dvireikšmių kintamųjų asociacijos matas, artimiausias standartiniams koreliacijų koeficientams yra polichorinės ir tetrachorinės koreliacijos. Todėl faktorių išskyrimo kategoriniams duomenims problemą darbe pasiūlyta spręsti Pearson‘o koreliacijų matricą pakeičiant polichorinių ir tetrachorinių koreliacijų matrica.

Tad, apibendrinant, darbo tikslui - Italijos smulkiųjų verslininkų pasitenkinimo darbu veiksmų analizei- įgyvendinti iškelti šie pagrindiniai uždaviniai: 1) geriausiai turimiems kokybiniais duomenims tinkančio modelio paieška; 2) kintamųjų modeliui parinkimas, vengiant multikolinearumo problemos; 3) kintamųjų skaičiaus sumažinimas faktorinės analizės pagalba; 4) pasirinkto modelio parametrų įvertinimas ir interpretacija atskirai vyrams ir moterims.

I. TEORINIS-METODINIS SKYRIUS

Šis skyrius skirtas trumpai teorijos, reikalingos verslininkų pasitenkinimo darbu daugiamatei analizei atlikti, apžvalgai.

Visų pirma bus supažindinta su pagrindinėmis darbe naudojamomis sąvokomis. Antroji skyriaus dalis skirta faktorinei analizei apžvelgti. Toliau trumpai apžvelgsime Apibendrintąjį Tiesinį modelį (Generalized Linear model) ir šiek tiek smulkiau panagrinėsime Logistinę, Daugiareikšmę logit bei Ranginę regresijas .

I.1. PAGRINDINĖS SĄVOKOS

I.1.1. Kategoriniai kintamieji

Kategoriniais arba kokybiniais vadinami kintamieji, kurių reikšmė yra atsakymas iš kategorijų aibės. Pagal matavimų skalę jie yra skirstomi į a) nominaliuosius ar vardinius, b) ranginius ir c) ,intervalinius‘.

Nominaliujų kintamųjų kategorijos nėra tarpusavyje palyginamos, jų reikšmės yra iš pavadinimų skalės, pvz. tautybė.

Ranginių kintamųjų kategorijas pagal reikšmę galima išrikiuoti, jos yra tarpusavyje palyginamos, pvz. išsilavinimas.

Kategoriniai ,intervaliniai‘ kintamieji turi kategorijas, kurios, kaip sako pavadinimas, yra sukurtos iš kiekybinės skalės intervalų, pvz. amžiaus grupės. Šių kintamųjų, kaip ir ranginių, kategorijas galime tarpusavyje palyginti ir išrikiuoti.

I.1.2. Optimalusis skalės transformavimas

Optimalusis skalės transformavimas (Optimal scaling) – tai daugiamatės analizės technika, kurią būtų galima apibrėžti kaip kiekybinių reikšmių priskyrimą kategoriniams duomenims, atliekant ,optimalųjį‘ skalės transformavimą.

Yra keletas Optimaliojo skalės transformavimo (OS) metodų, kurių bene plačiausiai naudojamas yra *regresija su OS*. Šio metodo principas yra kiekybinių reikšmių priskyrimas kategoriniams kintamiesiems tuo pat metu, kai mažiausių kvadratų principu vertinami modelio struktūriniai parametrai. Vienintelis apribojimas- kintamųjų monotoniškumas, o kiekybinės reikšmės parenkamos taip, kad geriausiai tiktų tiesinei regresijai (būtų maksimizuotas R^2).

Pritaikius OS galima analizuoti kartu bet kokios skalės kintamuosius, tačiau reikia atkreipti dėmesį, kad gautos kategorinių kintamųjų transformacijos prasmingos tik tai regresijai, su kuria jos gautos. Šis metodas įdiegtas SPSS programoje (CATREG algoritmas)

I.1.3. Multikolinearumas

Multikolinearumas- tai stiprios priklausomybės buvimas tarp keleto paaiškinamųjų kintamųjų. Jei priklausomybė sieja tik du paaiškinamuosius kintamuosius, tokia situacija vadinama *kolinearumu*.

Kai kintamieji multikolinearūs, sunku atskirti koreliuojančių kintamųjų įtaką prognozuojamai reikšmei, todėl multikolinearumo pasekmė gali būti modelio koeficientų ‚nepatikimumas‘ tiek dydžio ar ženklų, tiek stabilumo atžvilgiu.

Multikolinearumą galima pastebėti analizuojant koreliacijos bei dalinės koreliacijos koeficientus, tačiau šios statistikos ne visada jį išryškina. Todėl dažniausiai skaičiuojamas *dispersijos mažėjimo daugiklis VIF*.

Tarkim, r_j^2 yra regresijos modelio, kuriame X_j yra priklausomasis kintamasis, o $X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_n$ - nepriklausomieji kintamieji, determinacijos koeficientas. Tuomet kintamojo dispersijos mažėjimo daugiklis :

$$VIF = \frac{1}{1 - r_j^2}.$$

Laikoma, kad kintamasis yra multikolinearus, kai $VIF > 4$.

I.1.4. Polichorinės ir tetrachorinės koreliacijos

Tetrachorinės (polichorinės) koreliacijos- tai dvireikšmių (ranginių) kintamųjų asociatyvumo matas, besiremiantis šiomis prielaidomis:

- 1) už kategorinio kintamojo ‚slypi‘ tolydus latentinis (nestebimas) kintamasis;
- 2) jungtinis latentinių kintamųjų skirstinys yra bent apytikriai dvimatis normalusis.

Iš esmės tetrachorinės (polichorinės) koreliacijos yra latentinių tolydžiųjų kintamųjų Pearson‘o koreliacijos aproksimacijos.

Tetrachorinės koreliacijos koeficientą apibrėšiu pasinaudodama dažnių (contingency) lentele. Polichorinė koreliacija yra tiesiog tetrachorinės koreliacijos išplėtimas aukštesnės eilės dažnių lentelėms.

Tarkim, turim 2×2 *santykinų dažnių* lentelę :

		2-as kintamasis		
		1	2	
1-as kintamasis	1	a	b	a+b
	2	c	d	c+d
		a+c	b+d	1

čia a , b , c ir d yra poriniai santykiniai dažniai, gauti kintamųjų porinių stebėjimų dažnius padalijus iš bendro stebėjimų skaičiaus.

Pažymėkim $P_1 = a+b$ ir $P_2 = a+c$. Tai yra santykiniai 1-ojo ir 2-ojo kintamojo pirmosios kategorijos dažniai, tuo tarpu a yra jungtinis abiejų kintamųjų pirmosios kategorijos santykinis dažnis.

Pagal 1-ąją tetrachorinės koreliacijos prielaidą, už abiejų dvireikšmių kintamųjų, slypi nestebimi (latentiniai) tolydieji kintamieji X_1 ir X_2 , o tolydžiųjų kintamųjų suskirstymas į kategorijas vyksta nustačius slenkstines reikšmes (threshold) x_1 ir x_2 bei stebint, ar tolydieji kintamieji jas peržengė. Tarkim, 1-ajam kintamajam bus priskirta 2-oji kategorija, jei $X_1 > x_1$ kažkokiais šio kintamojo slenkstinei reikšmei x_1 .

Bet kokį tolydųjį atsitiktinį kintamąjį X galima transformuoti į standartinį normalųjį kintamąjį Z_x pagal šią formulę:

$$Z_x = \Phi^{-1}(\Phi_x(X)), \quad (1)$$

čia Φ_x - X sukauptoji pasiskirstymo funkcija (s. p. f.);

Φ - standartinio normaliojo skirstinio s. p. f. .

Kintamasis Z_x vadinamas X atitinkančiu standartiniu normaliuoju nukrypimu (deviate).

Transformacija (1) yra tiesinė ir monotonišė, todėl X_1 ir X_2 bei juos atitinkančių standartinių normalių nukrypimų koreliacijos koeficientai turėtų būti apytiksliai lygūs. Skaičiuojant transformuotųjų kintamųjų koreliacijos koeficientą reikėtų padaryti prielaidą, kad vektorius (Z_1, Z_2) turi dvimatį normalųjį skirstinį.

Taigi tarkim, $z_1 = \Phi^{-1}(P_1)$, $z_2 = \Phi^{-1}(P_2)$ yra standartiniai normalieji nukrypimai, atitinkantys santykinus 1-ojo ir 2-ojo kintamojo pirmosios kategorijos dažnius (arba ribines šios kategorijos tikimybes atitinkamiems kintamiesiems). Tuomet tetrachorinės koreliacijos koeficientas r yra toks, kad tenkinama:

$$a = \int_{z_1}^{\infty} \int_{z_2}^{\infty} \phi(x_1, x_2, r) dx_1 dx_2,$$

čia $\phi(x_1, x_2, r)$ yra dvimatė normalioji tikimybės pasiskirstymo funkcija:

$$\phi(x_1, x_2, r) = \frac{1}{2\pi\sqrt{1-r^2}} \exp\left[-\frac{1}{2(1-r^2)}(x_1^2 - 2rx_1x_2 + x_2^2)\right].$$

Geometriškai tiesės $x_1 = z_1$ ir $x_2 = z_2$ dalija šią dvimatę normaliąją pasiskirstymo funkciją į keturis kvadrantus, kurių tikimybės atitinka 2×2 dažnių lentelės santykinus porinius dažnius.

Ta pati logika taikoma ir polichorinės koreliacijos skaičiavimui, t.y. skaičiuojama koreliacija tarp standartinių normalių nukrypimų z_{1i} ir z_{2j} (čia $i=1, \dots, K$, kur K yra

I.2.2. Duomenų tikimo faktorinei analizei matai

Faktorinės analizės principas- kintamųjų skirstymas į grupes pagal jų panašumus, todėl išskiriant faktorius svarbu nustatyti, ar stebimi kintamieji tarpusavyje koreliuoja.

Vienas iš būdų- *Bartleto sferiškumo kriterijus*, padedantis nustatyti, ar iš viso tarp stebimų kintamųjų yra reikšmingai koreliuojančių. Juo naudojantis tikrinama hipotezė, kad koreliacijų matrica yra vienetinė, t.y. visi stebimi kintamieji nekoreliuoti. Todėl, jei taikant Bartleto sferiškumo kriterijų p-reikšmė $p \geq \alpha$, tai turimiems duomenims faktorinė analizė netaikytina, čia α - pasirinktasis reikšmingumo lygmuo.

Kitas duomenų tinkamumo faktorinei analizei matas- *Kaizerio-Mejerio-Olkino(KMO)*. Tai empirinių koreliacijos koeficientų didumų ir dalinių koreliacijos koeficientų didumų palyginamasis indeksas. Jis skaičiuojamas tokiu būdu:

$$KMO = \frac{\sum \sum_{i \neq j} r_{ij}}{\sum \sum_{i \neq j} r_{ij} + \sum \sum_{i \neq j} \tilde{r}_{ij}};$$

čia r_{ij} - kintamųjų X_i ir X_j koreliacijos koeficientas; \tilde{r}_{ij} - X_i ir X_j dalinės koreliacijos koeficientas. Laikoma, kad kai $KMO < 0,5$, faktorinė analizė nepriimtina.

Kiekvieno kintamojo stebėjimų tinkamumo matas, vadinamas *Kaizerio* matu :

$$MSA_i = \frac{\sum_{j \neq i} r_{ij}}{\sum_{j \neq i} r_{ij} + \sum_{j \neq i} \tilde{r}_{ij}}.$$

I.2.3. Faktorių išskyrimas pagrindinių komponentių analizės metodu

Vienas iš dažniausiai naudojamų faktorių išskyrimo būdų- pagrindinių komponentių analizė, kurios metu randamos tarpusavyje nekoreliuojančios kintamųjų X_1, \dots, X_k tiesinės daugdaros Y_1, \dots, Y_k , dar vadinamos *pagrindinėmis komponentėmis*.

Taigi pagrindinių komponentių analizės esmė- ieškomos pradinių kintamųjų tiesinės daugdaros :

$$Y_1 = \sum_{j=1}^k \alpha_{1j} X_j, \quad \dots, \quad Y_k = \sum_{j=1}^k \alpha_{kj} X_j,$$

tenkinančios sąlygas :

- 1) $\text{cov}(Y_i, Y_j) = 0, \quad i, j = 1, \dots, k, \quad i \neq j;$
- 2) $\mathbf{D}Y_1 \geq \mathbf{D}Y_2 \geq \dots \geq \mathbf{D}Y_k;$
- 3) $\sum_{i=1}^k \mathbf{D}Y_i = \sum_{i=1}^k \mathbf{D}X_i.$

Pagrindinių komponenčių radimo uždaviniu spręsti dažniausiai panaudojamas Lagranžo daugiklių metodas.

Išskirdami faktorius pasinaudosime šia pradinių kintamųjų išreiškimo pagrindinėmis komponentėmis formule:

$$X_i = \sum_{j=1}^k \alpha_{ji} Y_j, \quad i = 1, \dots, k, \quad j = 1, \dots, m.$$

Iš pradžių apskaičiuojami k pagrindinių komponenčių įverčiai :

$$\hat{Y}_i = \sum_{j=1}^k a_{ij} X_j, \quad i = 1, \dots, k;$$

kur a_{ij} yra koeficientų α_{ij} empiriniai įverčiai.

Latentiniais bendraisiais faktoriais laikomos m pirmųjų pagrindinių komponenčių, normuotų standartiniais nuokrypiais, t.y. :

$$\hat{F}_j = \frac{\hat{Y}_j}{\sqrt{s^2(\hat{Y}_j)}}.$$

Turėdami pagrindines komponentes, randame ir pačius faktorius. Faktorių svorių įverčiai yra

$$\hat{\lambda}_{ij} = a_{ji} \sqrt{s^2(\hat{Y}_j)}, \quad i = 1, \dots, k, \quad j = 1, \dots, m.$$

Specifinių faktorių įverčiai randami tokia lygybe:

$$\hat{e}_i = \sum_{j=m+1}^k a_{ji} \hat{Y}_j, \quad i = 1, \dots, k.$$

Taigi gauname tokį įvertį:

$$\hat{X}_i = \sum_{j=1}^m \hat{\lambda}_{ij} \hat{F}_j + \hat{e}_i, \quad i = 1, \dots, k;$$

čia visi bendrieji faktoriai F_j pagal prielaidas turėtų turėti vienetines dispersijas ir tarpusavyje nekoreliuoti. Be to, jie turėtų nekoreliuoti ir su specifiniais faktoriais.

Faktorius F_j laikomas susijusiu su tais kintamaisiais X_1, \dots, X_k , kuriems svorių įverčiai $\hat{\lambda}_{1j}, \dots, \hat{\lambda}_{kj}$ absoliučioju didumu ne mažesni kaip 0,4.

Pasitaiko, kad identifikuoti faktorius būna pakankamai sudėtinga, nes tas pats kintamasis būna susijęs su keliais faktoriais. Tuomet interpretacijos tikslais naudinga atlikti ortogonaliąją pradinių faktorių transformaciją, dar kitaip vadinamą *ortogonalioju sukimu*. Vienas iš plačiausiai naudojamų ortogonaliojų sukimų- VARIMAX.

Faktorių reikšmėms apskaičiuoti galima naudoti *mažiausių kvadratų procedūrą*. Tuo atveju gaunamos tokios formulės:

$$\hat{F}_j = \frac{\hat{Y}_j}{\sqrt{\lambda_j}};$$

čia \hat{Y}_j - j-osios pagrindinės komponentės įvertis, λ_j - jo tikrinė reikšmė.

Kitas būdas- *regresinės analizės technika*. Tarkim, faktoriai yra priklausomieji kintamieji, o pradiniai kintamieji X_i - nepriklausomi. Tada galime naudotis tokia lygtimi:

$$\hat{F}_j = \sum_{i=1}^k b_{ij} z_i, \quad j = 1, \dots, m;$$

čia \hat{F}_j yra j – ojo faktoriaus reikšmė, z_i - i-ojo kintamojo standartizuota reikšmė, b_{ij} - regresijos koeficientų įverčiai.

I.3. APIBENDRINTASIS TIESINIS MODELIS

Apibendrintasis Tiesinis modelis- tai Bendrojo Tiesinio modelio išplėtimas, kurio dėka galime įvertinti modelius ir tuo atveju, kai priklausomojo kintamojo pasiskirstymo funkcija nėra normalioji ar duomenys yra heteroskedastiški. Šio modelio parametrai β paprastai vertinami Didžiausio Tikėtimumo metodu, Kvazi-Didžiausio Tikėtimumo arba Bajeso metodais.

Modelio prielaida- priklausomasis kintamasis Y turi tam tikrą pasiskirstymo funkciją iš *eksponentinės šeimos*, kurią sudaro tokios pasiskirstymo funkcijos kaip Normalioji, Binominė, Puasono, Multinominė, Gamma, Neigiamoji Binominė ir kitos ir galioja :

$$E(Y) = \mu = g^{-1}(X\beta),$$

kur $X\beta$ - *tiesinis prediktorius*, t.y. tiesinė nežinomų parametų β kombinacija, o g vadinama *sąryšio funkcija* (link function).

Dispersiją paprastai galima išreikšti kaip vidurkio funkciją:

$$Var(Y) = V(\mu) = V(g^{-1}(X\beta)).$$

I.3.1. Apibendrintojo Tiesinio modelio elementai

Apibendrintąjį Tiesinį modelį (ATM) sudaro šie trys *elementai* :

- 1) pasiskirstymo funkcija f iš eksponentinės šeimos,
- 2) sisteminis elementas - tiesinis prediktorius $\eta = X\beta$,
- 3) sąryšio funkcija g tokia, kad $E(y) = \mu = g^{-1}(\eta)$.

Eksponentinė skirstinių šeima- tai tokie tikimybiniai skirstiniai su parametrais θ ir τ , kurių tankio (tolydžiajam Y skirstiniui) arba tikimybės masės (diskrečiajam skirstiniui) funkcijos gali būti išreikštos tokia forma :

$$f_y(y; \theta, \tau) = \exp\left(\frac{a(y)b(\theta) + c(\theta)}{h(\tau)} + d(y, \tau)\right).$$

čia τ – sklaidos (dispersion) parametras, kuris paprastai yra žinomas, o $a(\cdot)$, $b(\cdot)$ ir $c(\cdot)$ yra tam tikros žinomos funkcijos, besiskiriančios tarp skirtingų skirstinių.

Jei a - tapačioji (identity) funkcija, tuomet sakoma, kad skirstinys yra *kanoninėje formoje*. Jei ir parametras b - tapačioji funkcija, tuomet θ vadinamas *kanoniniu parametru*.

Sąryšio funkcija- tai ATM elementas, apibrėžiantis tiesinio prediktoriaus ir pasiskirstymo funkcijos sąryšį (per Y tikimybinių skirstinio vidurkį μ). Tarkim, $g(\mu_i)$ yra diferencijuojama monotoniška μ_i funkcija. Tuomet sąryšio funkciją galime apibrėžti tokiu būdu :

$$g(\mu_i) = \eta_i = \sum_{j=1}^k \beta_j x_{ij}, \quad i = 1, \dots, n.$$

Taigi sąryšio funkcija iš esmės nusako, koks ryšys sieja aiškinamuosius kintamuosius ir priklausomojo kintamojo vidurkį ir ,leidžia‘ , kad šis ryšys būtų netiesinis.

Šioje lentelėje yra pateiktos pagrindinės sąryšio funkcijos:

PASISKIRSTYMAS	PAVADINIMAS	SĄRYŠIO FUNKCIJA	VIDURKIO FUNKCIJA
Normalusis	<i>Identity</i> (Tapačioji)	$X\beta = \mu$	$\mu = X\beta$
Eksponentinis Gamma	<i>Inverse</i> (Atvirkštinė)	$X\beta = \mu^{-1}$	$\mu = (X\beta)^{-1}$
Puasono	<i>Log</i>	$X\beta = \ln(\mu)$	$\mu = \exp(X\beta)$
Binominis Multinominis	<i>Logit</i>	$X\beta = \ln\left(\frac{\mu}{1-\mu}\right)$	$\mu = \left(\frac{\exp(X\beta)}{1 + \exp(X\beta)}\right)$
	<i>Cloglog</i>	$X\beta = \ln(-\ln(1-\mu))$	$\mu = 1 - \exp(-\exp(X\beta))$

I.3.2. Logistinė regresija

Logistinės regresijos modelis- Apibendrintasis Tiesinis modelis su logit sąryšio funkcija. Jis naudojamas, kai turimas dvireikšmis (dar vadinamas Bernuliniu) priklausomasis kintamasis. Tokio priklausomojo kintamojo tikimybinį skirstinį galime nusakyti ‚sėkmės‘ tikimybe, t.y. $P(Y = 1) = \pi$ ir ‚nesėkmės‘ tikimybe $P(Y = 0) = (1 - \pi)$, o skirstinio vidurkis $E(Y) = \pi$ ir dispersija $\text{Var}(Y) = \pi(1 - \pi)$.

Modelį galime užrašyti tokiu būdu:

$$\text{logit}(\pi(x)) = \beta_0 + \sum_{j=1}^k \beta_j x_j = X\beta,$$

kur $\text{logit}(\pi(x))$ yra natūralusis tikimybių santykio (odds ratio) logaritmas :

$$\text{logit}(\pi(x)) = \log \left[\frac{\pi(x)}{1 - \pi(x)} \right],$$

o $\pi(x)$ yra tikimybė, kad Y įgis reikšmę 1 su aiškinamųjų kintamųjų funkcija. Šią tikimybę dar galime užrašyti tokiu būdu :

$$\pi(x) = \frac{\exp(X\beta)}{1 + \exp(X\beta)}.$$

Modelio parametrai įvertinami Didžiausio tikėtinumo metodu.

I.3.2.1. Statistinės išvados

- **Modelio X^2 suderinamumo kriterijus**

Šis kriterijus skirtas patikrinti, ar apskritai modelyje yra reikšmingų kintamųjų, t.y. ar nors vienas $\beta_j \neq 0$. Nulinė hipotezė- $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$. Kriterijaus statistiką apskaičiuojame tokiu būdu :

$$X^2 = -2 \ln L(\hat{\beta}(M_0)) + 2 \ln L(\hat{\beta}(M)),$$

čia $L(\hat{\beta}(M_0))$ - Didžiausio tikėtinumo funkcijos įvertis modelio, turinčio tik konstantą,

o $L(\hat{\beta}(M))$ - Didžiausio tikėtinumo funkcijos įvertis analizuojamam modeliui.

Pasirinktam reikšmingumo lygmeniui α hipotezė H_0 atmetama, jei $X^2 > X_{\alpha}^2(k)$, kur $X_{\alpha}^2(k)$ yra X^2 skirstinio su k laisvės laipsnių α lygmens kritinė reikšmė.

- **Voldo (Wald) kriterijus**

Tai Studento kriterijaus tiesinėje regresijoje analogas. Tikrinama nulinė hipotezė, ar konkretus koeficientas nelygus nuliui, t.y. $H_0: \beta_j = 0$ prieš $H_1: \beta_j \neq 0$.

Wald statistika $W = (\hat{\beta}_j / SE_j)^2$, kur SE_j - $\hat{\beta}_j$ standartinio nuokrypio įvertis. Ši statistika turi χ^2 skirstinį su 1 laisvės laipsniu.

- **Determinacijos koeficientai**

Kaip ir tiesinėje regresijoje, naudojami įvertinti, ar nepriklausomų kintamųjų ir Y priklausomybė yra didelė. Yra keletas tradicinio R^2 analogų Logistinei regresijai:

Makfadeno pseudodeterminacijos koeficientas:

$$r_M^2 = 1 - \frac{-2 \ln L(\hat{\beta}(M))}{-2 \ln L(\hat{\beta}(M_0))} = 1 - \frac{\ln L(\hat{\beta}(M))}{\ln L(\hat{\beta}(M_0))},$$

kur $L(\hat{\beta}(M_0))$ - D. T. funkcijos įvertis modeliui tik su konstanta;

$L(\hat{\beta}(M))$ - D. T. funkcijos įvertis analizuojamam modeliui.

Kokso-Snelo pseudodeterminacijos koeficientas:

$$r_{CS}^2 = 1 - \left(\frac{L(\hat{\beta}(M_0))}{L(\hat{\beta}(M))} \right)^{2/n},$$

kur n - stebėjimų skaičius.

Šis determinacijos koeficientas mažesnis už 1.

Nagelkerkė pseudodeterminacijos koeficientas:

Tai normuotasis Kokso-Snelo pseudodeterminacijos koeficientas. Jo formulė:

$$r_N^2 = \frac{r_{CS}^2}{\max r_{CS}^2} = \frac{r_{CS}^2}{1 - (L(\hat{\beta}(M_0)))^{2/n}}.$$

I.3.3. Daugiareikšmis logit modelis

Daugiareikšmis(multinomial) logit modelis- tai logistinės regresijos išplėtimas, kai priklausomasis kintamasis gali įgyti daugiau negu 2 reikšmes ir tos reikšmės yra vardinės. Kartais naudojamas ir ranginiam priklausomam kintamajam, kai pažeidžiama lygiagrečiųjų regresijos tiesių (parallel regression lines) prielaida, reikalinga ranginei regresijai įvertinti.

Šio modelio idėja yra dviejų priklausomojo kintamojo kategorijų lyginimas vienu metu, kur kaip bazinė gali būti pasirinkta bet kuri priklausomojo kintamojo reikšmė, o pagrindinė prielaida- nepriklausomumas nuo nesvarbių alternatyvų (Independence of

Irrelevant Alternatives). Ši prielaida reiškia, kad bet kurių dviejų kategorijų (alternatyvų) tikimybių santykis (odds) nepriklauso nuo kitų alternatyvų, t.y. jei į modelį įtrauktume daugiau alternatyvų, arba kurią nors jų pašalintume iš modelio, bet kurių dviejų kategorijų tikimybių santykis nepasikeistų.

Modelį galima užrašyti tokiomis formulėmis:

$$P(y_i = j) = \frac{\exp(X_i \beta_j)}{\sum_{j=1}^J \exp(X_i \beta_j)} \quad \text{ir} \quad P(y_i = 0) = \frac{1}{\sum_{j=1}^J \exp(X_i \beta_j)},$$

čia y_i - priklausomojo kintamojo stebėjimai, X - aiškinamųjų kintamųjų vektorius, β_j - nežinomi parametrai, $j=0,..J$ – galimos alternatyvos, kur 0 – bazinė alternatyva.

Alternatyvų tikimybių suma lygi vienetui.

Daugiareikšmio logit modelio parametrai paprastai vertinami Didžiausiojo tikėtino metodo, o modelio tikimo duomenims kriterijai bei determinacijos koeficientai apskaičiuojami taip pat, kaip Logistinės regresijos atveju.

I.3.4. Ranginė (ordinal) regresija

Ranginė regresija- modelis, naudojamas tuo atveju, kai priklausomasis kintamasis įgyja reikšmes, kurias galima surikiuoti.. Jis, kaip ir ankstesni aptarti modeliai, yra iš Apibendrintojo tiesinio modelio šeimos, o priklausomojo kintamojo tikimybes galima modeliuoti keletu sąryšio funkcijų, kurių plačiausiai naudojamos yra logit ir cloglog, kurias vėliau aptarsime plačiau.

Bendru atveju ranginę regresiją galime išivaizduoti, kaip modelį su nestebimu (latentiniu) tolydžiuoju priklausomu kintamuoju $Y_i^* \sim \text{Logit}(y_i^* | \mu_i)$. Stebimas priklausomasis kintamasis Y_i reikšmes įgyja tokiu principu :

$$Y_i = j, \text{ jeigu } \gamma_{j-1} \leq Y_i^* \leq \gamma_j, \text{ kai } j = 1, \dots, J,$$

čia γ_l ($l = 0, \dots, J$) yra slenkstiniai (threshold) parametrai tokie, kad $\gamma_l < \gamma_m$ visiems $l < m$ ir $\gamma_0 = -\infty$, o $\gamma_J = \infty$.

I.3.4.1. Lygiagrečiųjų tiesių (plokštumu) prielaida

Svarbiausia modelio prielaida- *lygiagrečiųjų tiesių (plokštumu) prielaida*. Ji reiškia, kad priklausomojo kintamojo ir aiškinamųjų kintamųjų sąryšis yra nepriklausomas nuo kategorijos- regresijos koeficientai vienodi visoms kategorijoms, o skiriasi tik skirtingų kategorijų konstantos. Modelio konstantas galime interpretuoti kaip slenkstines reikšmes, pagal kurias tolydusis latentinis kintamasis suskirstytas į stebimas kategorijas. Geometriškai ši prielaida reiškia, kad rezultatai (įvertintos tikimybės) yra lygiagrečių tiesių (ar plokštumu) aibė- kiekvienai kategorijai po vieną.

- **Prielaids tikrinimas**

Tarkim, priklausomasis kintamasis įgyja reikšmes $1, \dots, k, k+1$, kur $(k+1) > 2$ ir turim s aiškinamųjų kintamųjų. Nagrinėkime bendrą sukauptąjį (cumulative) modelį be lygiagrečiųjų tiesių prielaidos :

$$g(P(Y \leq i | x)) = (1, x')\theta_i, \quad 1 \leq i \leq k,$$

čia $g(\cdot)$ yra sąryšio funkcija, o $\theta_i = (\alpha_i, \beta_{i1}, \dots, \beta_{is})'$ yra nežinomų parametru vektorius, kurį sudaro konstantos (intercept) α_i ir s krypties parametru $\beta_{i1}, \dots, \beta_{is}$. Parametru vektorius šiam bendrajam sukauptajam modeliui yra $\theta = (\theta_1', \dots, \theta_k')$.

Pagal nulinę lygiagretumo hipotezę $H_0 : \beta_{1m} = \beta_{2m} = \dots = \beta_{km}$, $1 \leq m \leq s$ yra vienintelis bendras krypties parametras kiekvienam iš s aiškinamųjų kintamųjų. Tegul β_1, \dots, β_s yra bendri krypties parametrai. Tegul $\hat{\alpha}_1, \dots, \hat{\alpha}_k$ ir $\hat{\beta}_1, \dots, \hat{\beta}_s$ yra parametru maksimalaus tikėtinumo įverčiai. Tuomet, pagal H_0 , maksimalaus tikėtinumo įvertis θ yra :

$$\hat{\theta}_0 = (\hat{\theta}_1', \dots, \hat{\theta}_k')', \quad \text{kur } \hat{\theta}_i = (\hat{\alpha}_i, \hat{\beta}_1, \dots, \hat{\beta}_s)'$$

o kriterijaus statistika $U'(\hat{\theta}_0)I^{-1}(\hat{\theta}_0)U(\hat{\theta}_0)$ turi asimptotinę X^2 skirstinį su $s(k-1)$ laisvės laipsnių. Jei nulinė hipotezė atmetama, vadinasi modelis duomenims netinka ir matyt reikėtų taikyti daugiareikšmį logit modelį.

Ranginės regresijos modelio parametru vertinimas ir modelio tikimo duomenims statistikos tokios, kaip ir Logistinės regresijos, kurias apibrėžiau kiek anksčiau..

Toliau panagrinėsime du populiariausius ranginės regresijos modelius – Proporcionaliųjų galimybiu modelį (logit sąryšio funkcija) ir Proporcionaliųjų riziku modelį (cloglog s.f.).

I.3.4.2. Proporcionaliųjų galimybiu modelis

Tai ranginės regresijos modelis su logit sąryšio funkcija, dar vadinamas Sukauptuoju Logit modeliu. Dažniausiai naudojamas, kai priklausomojo kintamojo kategorijos yra vienodai pasiskirsčiusios. Modelį per sąryšio funkciją galima užrašyti tokiu būdu:

$$f[g_j(X)] = \ln \left\{ \frac{g_j(X)}{1 - g_j(X)} \right\} = \ln \left\{ \frac{P(Y \leq y_j | X)}{P(Y > y_j | X)} \right\} = \alpha_j + \beta X,$$

čia $j=1, \dots, k-1$ yra priklausomojo kintamojo kategorijos, o

$$g_j(X) = \exp(\alpha_j + \beta X) / [1 + \exp(\alpha_j + \beta X)].$$

Tuo tarpu α_j yra slenkstinės reikšmės kiekvienai sukauptajai tikimybei.

I.3.4.3. Proporcionaliųjų rizikų modelis

Tai ranginės regresijos modelis su cloglog sąryšio funkcija, dar vadinamas Continuation Ratio modeliu. Paprastai naudojamas, kai labiau tikėtinos yra aukštesniojo rango priklausomojo kintamojo kategorijos. Modelį galim užrašyti taip :

$$f[g_j(X)] = \ln\{-\ln[1 - g_j(X)]\} = \ln\left\{-\ln\left[\frac{P(Y = y_j | X)}{P(Y > y_j | X)}\right]\right\} = \alpha_j + \beta X,$$

kur $j=1, \dots, k-1$ yra priklausomojo kintamojo kategorijos, o

$$g_j(X) = 1 - \exp\{-\exp(\alpha_j + \beta X)\}.$$

II. TIRIAMASIS-ANALITINIS SKYRIUS

Ši dalis skirta ištirti verslininkų pasitenkinimo darbu veiksniais pasinaudojant teorinėje dalyje aprašytais metodais. Pradėsime nuo duomenų apžvalgos, multikolinearumo tyrimo, vėliau bus atlikta faktorinė analizė, o gautus faktorius panaudosime kurdami modelius. Palyginę modelius pagal jų prielaidas, tikimo duomenims statistikas, prognozių efektyvumą išsirinksime tinkamiausią duomenims modelį ir galėsime panagrinėti, kaip įvairūs veiksniai įtakoja verslininkų pasitenkinimą darbu. Bus sukurti ir palyginti atskiri modeliai vyrams ir moterims ir padarytos svarbiausios išvados.

II.1. Duomenys

Šiame skyrelyje bus supažindinta su duomenų šaltiniu, kintamaisiais, jų transformacijomis ir bus atlikta pirminė duomenų apžvalga lyginant pagrindines jų statistikas tiek bendrai, tiek atskirai vyrams ir moterims. Taip pat bus atliktas multikolinearumo tyrimas ir pasirinktas metodas koreliuotų kintamųjų problemai išspręsti.

II.1.1. Duomenų šaltinis ir apklausos dalys

Duomenys, kuriais naudosimės tirdami verslininkų pasitenkinimą darbu yra iš Italijos Veneto srities verslininkų tyrimo, kurį 2006 metais pradėjo Padujos universiteto Statistikos departamentas. Šio tyrimo tikslas buvo susidaryti bendrą vaizdą apie Veneto regiono verslininkų padėtį, lūkesčius ir perspektyvas, ypač didelį dėmesį skiriant vyrų ir moterų palyginimui. Apklausa atlikta telefonu, naudojantis CATI (Computer Assisted Telephone Interviewing) metodu.

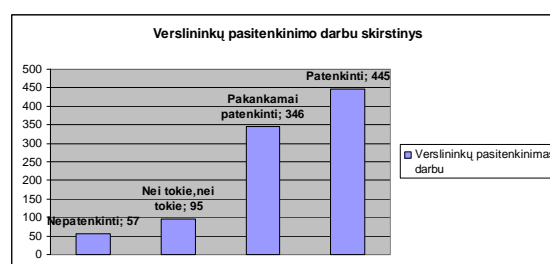
Imtį sudaro 1216 stebėjimų (826 moterų ir 390 vyrų) iš 113 000 Veneto regiono verslininkų populiacijos (26 000 moterų ir 77 000 vyrų).

Verslininkų apklausą (115 atskirų klausimų) sudarė šios dalys:

- asmens identifikavimas (lytis, gyvenamoji vieta, šeimyninė padėtis);
- išsilavinimas ir lankyti kursai;
- sutuoktinis ir šeima ;
- firmos savybės;
- veikla, susijusi su firma(darbu);
- laisvalaikis;
- lūkesčiai ir profesinės perspektyvos.

II.1.2. Kintamieji ir jų transformacijos

Mūsų tiriamasis kintamasis- verslininkų pasitenkinimas darbu, įvertintas 4 taškų Likert skalėje, kurio skirstinys pavaizduotas šiame pav. :



Lyginant pasitenkinimo darbu pasiskirstymą tarp vyrų ir moterų (žr. Priedai p.), galima pastebėti, kad tiriamojo kintamojo rangai tarp abiejų lyčių pasiskirstę labai panašiai, o skirstinys yra asimetriškas- dominuoja aukštesni pasitenkinimo darbu rangai.

Renkantis aiškinamuosius kintamuosius, pirmiausiai buvo atliktas asociacijų/koreliacijų su priklausomuoju kintamuoju tyrimas ir tolimesnei analizei pasilikti tik tie kintamieji, kurie fiksuojant reikšmingumo lygmenį $\alpha=0,01$ buvo statistiškai reikšmingai susiję su pasitenkinimu darbu. Tad pradinį 114 kintamųjų skaičių pavyko sumažinti iki 35 potencialių aiškinamųjų kintamųjų. Prieš pateikiant duomenis daugiamatei analizei, buvo atliktos ir kai kurių kintamųjų transformacijos (kategorijų apjungimas arba kiekybinių kintamųjų suskirstymas intervalais) modelio supaprastinamo, interpretacijos palengvinimo ar sąryšio su priklausomuoju kintamuoju sustiprinimo tikslais. Visi kintamieji ir jų transformacijos aprašytos prieduose (p.30-33).

II.1.3. Trumpa imties apžvalga

Prieš pradėdant daugiamatę duomenų analizę verta susipažinti su pagrindinėmis kintamųjų aprašomosiomis statistikomis „bendram vaizdui“ apie imtį susidaryti.

Žvilgtelėję į dažnių lenteles (priedai, p. 30-33), galime pastebėti, kad dominuoja vidutinio ir vyresnio amžiaus verslininkai (jaunesni negu 35 m. sudaro tik 11 % imties), vyrauja susituokę, turintys vidutiniškai 1-2 vaikus bei aukštesnįjį arba vidurinį išsilavinimą individai.

Kalbant apie firmą, vidutinis darbuotojų skaičius yra apie 7-8, daugiau negu pusėje firmų yra iki 5 darbuotojų, taigi mūsų imtyje vyrauja smulkaus verslo atstovai. Daugiau negu pusė jų firmą įkūrė patys ir apie trečdalis naudojo nuosavą kapitalą firmos kūrimo procese. Firmos yra nedidelės, turbūt todėl ir lėšų investicijoms į informacinės sistemos plėtimą ar personalo tobulinimą skiria tik apie 6 procentus apklaustųjų firmų savininkų. Kokybės kontrolės patvirtinimas taip pat nesvarbus apie 80-ies procentų respondentų firmoms.

Daugumai verslininkų darbas sektoriuje buvo jų pasirinkimas (72%), o dominuojantis tapimo verslininku motyvas- noras pradėti savarankišką gyvenimą. Vyrauja pesimistiškos ateities perspektyvos- tik apie 23% verslininkų mano, kad situacija sektoriuje pagerės, o manančiųjų, kad pablogės yra beveik 44%.

Žvilgtelėjus į laiką, skiriamą darbui, matome, kad apie 40% respondentų dirba 10-12 valandų per dieną (iš vyrų net 52%) ir apie 9% dirba net >12 valandų per dieną. Vyrų atrodo linkę skirti daugiau laiko darbui negu moterys, virš 10 valandų per dieną darbui skiria 61% vyrų ir 43% moterų. Dauguma verslininkų sudaro darboholikų išpūdį - 62% dirba ir savaitgaliais, 61% laisvu laiku tvarko atsilikusius darbo reikalus, daugiau negu pusė apskritai neturi laisvo laiko, o vidutiniškai atostogauja 2 savaites per metus (17% iš viso neturi atostogų).

Lyginant moteris ir vyrus, labai didelių skirtumų kintamųjų aprašomosiose statistikose nepastebėta. Visgi lyginant su vyrais, galima pastebėti, kad imtį sudaro jaunesnės moterys, su aukštesniu negu vyrai išsilavinimu, didesnė jų dalis kūrė firmą siekdamas savarankiško gyvenimo pradėjimo, tačiau mažesnė dalis negu vyrų tarpe dirbti sektoriuje pasirinko pačios. Be to, kaip jau buvau užsiminusi, moterys linkusios mažiau laiko negu vyrai skirti darbui.

II.1.4. Multikolinearumo tyrimas

Multikolinearumas- dažnai pasitaikantis reiškinys didelėje aiškinamųjų kintamųjų grupėje, ypač tarp ranginių ir dvireikšmių kintamųjų. Jo galimi padariniai – aukštos standartinės paklaidos, dėl kurių svarbūs kintamieji tampa nereikšmingi, regresijos koeficientų dydis ar ženklas, neatitinkantis realybės bei koeficientų nestabilumas. Todėl siekiant išvengti šio nepageidaujamo reiškinio, reikėtų ištirti, kurie kintamieji yra stipriai susiję ir priimti atitinkamus sprendimus multikolinearių kintamųjų elgesio atžvilgiu. Populiariausi multikolinearumui identifikuoti naudojami metodai yra koreliacijos, dalinės koreliacijos, Dispersijos mažėjimo daugiklis (VIF) arba Tolerancija= $1/VIF$.

Atlikę susijusių kintamųjų analizę (prieduose p. 34 yra pateikti visų aiškinamųjų kintamųjų VIF, Tolerancijos bei 1-ojo ir 2-ojo tipo dalinės koreliacijos koeficientai), matome, kad nemaža dalis kintamųjų yra stipriai susiję – net 14-ai iš 34 kintamųjų VIF bei Tolerancijos statistikos rodo „per didelį multikolinearumą“. Panagrinėjus atidžiau, galima išskirti penkias labiausiai tarpusavyje susijusių kintamųjų grupes, kurių viduje reikėtų ieškoti „nepageidajamų“ kintamųjų (šios grupės išskirtos atsižvelgiant į statistiškai reikšmingiausias koreliacijas - prieduose p. 35-37 pateikti stipriausiai susijusių kintamųjų Spearman'o koreliacijos koeficientai).

Išanalizavus kintamųjų tarpusavio priklausomybę, multikolinearių (arba tiesiog kolinearių) aiškinamųjų kintamųjų problemą nuspręsta eliminuoti tokiu būdu :

- 1) atsisakyti *šeimos narių, vaikų* bei *vaikų gyvenančių kartu* skaičių atspindinčių kintamųjų- iš stipriai koreliuotų pasiliktas tik kintamasis „*turi mokyklinio amžiaus vaikų*“ - jo asociacija su priklausomuoju kintamuoju didžiausia (Spearman'o Rho yra 0,287 lyginant su atitinkamai 0,093, 0,172 ir 0,188 iš analizės „pašalintiems“ kintamiesiems);
- 2) 2-ojoje grupėje stipriausia priklausomybė siejo kintamuosius „*firmą įkūrė pats/pati*“ ir „*firmai kurti naudotas nuosavas kapitalas*“, iš kurių pastarojo nuspręsta atsisakyti, nes kintamasis „*firmą įkūrė pats/pati*“ yra labiau susijęs su pasitenkinimu darbu (Spearman'o Rho yra 0,1875 lyginant su 0,1003);
- 3) 3-iosios grupės labiausiai susijusius kintamuosius, t.y. „*investuoja į personalo kvalifikacijos kėlimą*“ ir „*investuoja į firmos informacinės sistemos tobulinimą*“ nuspręsta transformuoti- sudaryti naują kintamąjį apjungimo būdu (naujo kintamojo kategorijos ir dažniai prieduose p. 37) ;
- 4) iš dviejų labiausiai susijusių 4-osios grupės kintamųjų- „*tenka keliauti darbo reikalais*“ ir „*dienų komandiruotėje skaičius*“ nuspręsta pasilikti pirmąjį – verslininkų, keliaujančių darbo reikalais mūsų imtyje yra nedaug (12 %), tad turi didesnę prasmę tirti pasitenkinimo darbu priklausomybę nuo komandiruočių buvimo negu nuo dienų komandiruotėse skaičiaus;
- 5) 5-ojoje grupėje daugiausiai problemų kėlė „*valandų skaičius poilsiui*“ (jo ir VIF didžiausias- 21) , susijęs su visais pastarosios grupės kintamaisiais. Tad jo nuspręsta atsisakyti.

Rezultatas- š 34 paaiškinamųjų kintamųjų tolesnei analizei liko 27.

Reikėtų pastebėti, kad multikolinearių kintamųjų problemos sprendimas yra labai subtilus dalykas ir kiekvienas tyrėjas gali rasti vis kitokių išeičių. Prieš priimdama kokį nors sprendimą, jo efektyvumą tikrinau ir empiriškai ir iš visų svarstytų galimybių šios išeitys atrodė pagrįščiausios ir geriausiai tinkančios modeliui.

II.2. Faktorinė analizė

Faktorinė analizė- daugiamačės statistikos metodas, taikomas siekiant pakeisti grupę koreliuojančių kintamųjų bendrųjų faktorių rinkiniu, kurie dažnai interpretuojami kaip latentiniai kintamieji. Tokiu būdu yra kondensuojama informacija, taip pat eliminuojama kintamųjų tarpusavio koreliacijos problema jai esant.

Atliekant multikolinearumo tyrimą buvo galima pastebėti, kad kintamieji yra tarsi susiskirstę į grupes pagal tarpusavio sąryšį. Ir nors eliminavus stipriausiai susijusius kintamuosius lyg ir išsprendėme multikolinearumo problemą, tačiau faktorinė analizė būtų labai naudinga modelio supaprastinimui- vietoj to, kad analizuotumėme atskirai kiekvieno kintamojo įtaką pasitenkinimui darbu, galbūt mums pavyktų išskirti kažkokius bendrus veiksnius, atspindinčius grupę kintamųjų, nuo kurių priklauso verslininkų pasitenkinimas darbu.

Tačiau norėdami taikyti faktorinę analizę susiduriame su problema- viena iš modelio prielaidų- stebimų koeficientų normalusis skirstinys, nes išskiriant faktorius remiamasi Pearson'o koreliacijų arba kovariacijų koeficientais kintamųjų priklausomumui nustatyti. Kaip žinia, dauguma mūsų kintamųjų yra ranginiai arba dvireikšmiai, todėl Pearson'o koreliacijos koeficientai pateiktų iškreiptas reikšmes ir rezultatais nevertėtų pasitikėti.

Kategorinių kintamųjų sąryšiui apskaičiuoti naudojami įvairūs asociacijos matai, tokie kaip Kendal , Spearman koreliacijų ar konkordacijos koeficientai. Ranginių ar dvireikšmių kintamųjų priklausomumui įvertinti naudojamos ir polichorinės arba tetrachorinės koreliacijos, kurios yra laikomos Pearson'o koreliacijų aproksimacijomis. Jų taikymo prielaida- tarimas, jog stebimas ranginis ar dvireikšmis kintamasis yra sudarytas iš tolydaus latentinio kintamojo. Būtent šiais koreliacijos koeficientais ir pasinaudosime taikydami faktorinę analizę.

Siekiant išvengti nereikšmingų kintamųjų apjungimo į faktorius, visų pirma buvo sudarytas modelis iš likusiųjų analizei kintamųjų taikant backward, stepwise ir forward modelio sudarymo būdus. Visais šiais metodais buvo gautas tas pats modelis (priedai, p.), kur visi naudoti aiškinamieji kintamieji yra statistikai reikšmingi. Tad labiausiai susijusių ieškosime tarp visų kintamųjų, išskyrus vardinius ,tapimo verslininku motyvai' , šeimyninė padėtis' ir ,lytis', kuriems polichorinės ar tetrachorinės koreliacijos neturi prasmės.

Apskaičiavus polichorinių bei tetrachorinių koreliacijų koeficientus, buvo atrinkti labiausiai susiję kintamieji (tokių rasta 12) ir jiems sudaryta polichorinių-tetrachorinių koreliacijų matrica. (priedai, p. 38-40) Ši matrica buvo pateikta faktorinei analizei, pasirinkus faktorių išskyrimo pagrindinėmis komponentėmis metodą.

Visų pirma buvo patikrintas duomenų tikimas faktorinei analizei- tiek KMO, tiek Bartleto sferiškumo testai patvirtino, kad mūsų duomenys analizei tinka. Atsižvelgiant į prasmingumą , paaiškinamos dispersijos procentą, tikrinių reikšmių grafiką buvo išskirti 3 faktoriai, kartu paaiškinantys 51% dispersijos. Faktorių interpretacijai supaprastinti, o kartu ir siekiant garantuoti koreliacijų tarp faktorių nebuvimui, buvo atliktas ortogonalusis faktorių sukimas VARIMAX.

Analizuojant kintamuosius, labiausiai susijusius su faktoriais, pastariesiems buvo suteiktos ,laisvalaikio' , ,firmos' ir ,darbo laiko' interpretacijos. Faktorių reikšmėms apskaičiuoti naudotas regresijos metodas.

Rezultatų palyginimui buvo atlikta faktorinė analizė tiems patiems kintamiesiems naudojantis Pearson'o koreliacijomis (priedai, p.46-47) . Kaip ir galima buvo tikėtis, turimiems kintamiesiems šis būdas nelabai tiko arba ,tiksliau, faktorinė analizė su Pearson'o koreliacijomis buvo mažiau efektyvesnė už atliktą su polichorinėmis bei tetrachorinėmis koreliacijomis. Įprastu būdu išskirti faktoriai paaiškina mažiau dispersijos- pirmieji trys paaiškina tik 35% variacijos lyginant su 51% , kai naudotos polichorinės koreliacijos. Norėdami gauti tą patį paaiškintos dispersijos procentą, turėtumėme iš 12 kintamųjų išskirti net 5 faktorius.

Tad galime padaryti išvadą, kad norėdami taikyti faktorinę analizę ranginiams arba dvireikšmiams kintamiesiems, turėtumėme naudotis ne Pearson'o koreliacijų koeficientais, o koku nors alternatyviu asociacijų matu. Turbūt tinkamiausios šiam tikslui būtų polichorinės arba tetrachorinės koreliacijos, kurios yra laikomos Pearson'o koreliacijų aproksimacijomis.

II.3. Modelio parinkimas

Kaip jau buvo paminėta kiek anksčiau, turėdami ranginį priklausomąjį kintamąjį galime bandyti rinktis tarp keleto modelių: galime taikyti tiesinę regresiją su optimaliu skalės transformavimu, ranginę regresiją sudvireikšminus priklausomąjį kintamąjį, ranginę regresiją su keletu galimų sąryšio funkcijų (dažniausiai naudojamos logit ir cloglog) arba jei netinka ranginė regresija dėl proporcingųjų tiesių prielaidos netenkinimo, galima naudoti daugiareikšmę logistinę regresiją.

Regresijos su optimaliu skalės transformavimu esmė- kategorinių kintamųjų pavertimas į kiekybinius reikšmes parenkant taip, kad būtų maksimizuotas R^2 . Tokio tipo regresija geriau tinka kintamiesiems, turintiems daugiau kategorijų, todėl mūsų priklausomajam kintamajam turint jų tik keturias, tiesinė transformacija į kiekybinę skalę yra sunkiai įgyvendinama. Tai matosi ir iš rezultatų (priedai, p. 48-49)- reikšmingų kintamųjų dalis yra nedidelė, be to, optimizuojant skalę kategoriniams kintamiesiems yra pakankamai sudėtinga jų interpretacija. Todėl šio modelio atsisakyta pirmiausiai.

Kitas būdas- ranginio kintamojo pavertimas dvireikšmiu ir taikoma logistinė regresija. Šiuo atveju sumažindami kategorijų skaičių prarandame svarbios informacijos apie priklausomąjį kintamąjį, nors tai ir labai supaprastina pačią analizę. Neturint kitų būdų, padedančių išsaugoti daugiau informacijos, susijusios su priklausomuoju kintamuoju, logistinė regresija gali būti geriausia išeitis.

Visgi turime dar vieną alternatyvą- jei tenkinama lygiagrečiųjų tiesių prielaida, ranginiam kintamajam modeliuoti geriausias būdas gali būti ranginė regresija. Ji išsaugo visą informaciją apie priklausomąjį kintamąjį, be anksčiau paminėtos lygiagrečiųjų tiesių prielaidos, neturi jokių kitų apribojimų duomenų formai ir struktūrai. Be to, ranginė regresija turi keletą sąryšio funkcijų, geriau modeliuojančių vieną ar kitą priklausomojo kintamojo skirstinį. Taigi šis būdas ranginiams priklausomiesiems kintamiesiems modeliuoti bene pats geriausias, tačiau tik tuo atveju kai tenkinama lygiagrečiųjų tiesių prielaida.

Visų sukurtų modelių rezultatai pateikti prieduose (p. 48-49). Ranginei regresijai kurti naudojami dvejomis dažniausiai taikomomis sąryšio funkcijomis, t.y. logit ir cloglog. Nors nėra nustatytos taisyklės, kaip pasirinkti sąryšio funkciją, tačiau laikoma, kad logit yra tinkamesnė, kai kintamojo skirstinys yra gana simetriškas, o cloglog geriau naudoti

turint priklausomąjį kintamąjį, kurio skirstinio didžiausią dalį sudaro aukštesniojo rango kategorijos.

Kaip jau minėjau, taikant ranginę regresiją, labai svarbi yra lygiagrečiųjų tiesių prielaida. Mūsų duomenims, kaip matome iš rezultatų (priedai p. 50), ši prielaida tenkinama abiejoms sąryšio funkcijoms, todėl galime drąsiai analizuoti rezultatus.

Tirdami modelių koeficientus (paprastumo dėlei nagrinėsime tik jų ženklus, norėdami palyginti koeficientų dydžius turėtume naudotis jų galimybių santykiais (odds ratio), tačiau tai neapima šios analizės tikslų) matome, kad tiek logistinė, tiek ranginė regresija su logit bei cloglog sąryšio funkcijomis yra panašios koeficientų atžvilgiu. Dažniausiai skiriasi tik statistiškai nereikšmingų kintamųjų ženklai. Siekiant išsirinkti labiausiai duomenims tinkantį modelį, verta paanalizuoti determinacijos koeficientus ir prognozavimo efektyvumą. Rezultatai patvirtina, kad kaip ir buvo spėjama, mūsų duomenims iš nagrinėtų modelių labiausiai tinkanti yra ranginė regresija su cloglog sąryšio funkcija – aukščiausi yra tiek determinacijos koeficientai, tiek prognozių efektyvumo statistikos (priedai p. 50-51). Visgi skirtumas nėra itin ryškus, tad modelio išvados matyt labai nenukentėtų pasirinkus ir ranginį logit modelį.

Regresija su optimaliuoju skalės transformavimu neatrodo tinkama mūsų ranginiam priklausomajam kintamajam, o logistinės regresijos, nors jos rezultatai ir panašūs į gautuosius su rangine regresija, vertėtų atsisakyti – informacijos praradimas tokioje situacijoje yra bevertis.

Multinominę regresiją vertėtų naudoti tik tada, kai netenkinama lygiagrečiųjų tiesių ar tiesinių daugdarų prielaida, nes priešingu atveju modelio kiekvienai kategorijai atskirai tyrimas netenka prasmės. Analizės pradžioje buvo panaudotas ir šis metodas priklausomajam kintamajam tirti, bet jo interpretacija yra ir pakankamai komplikuota, ir neturi daug prasmės žinant, kad pasitenkinimas darbu yra ranginis ir savo „prigimtim“.

Pabaigai prieduose (p. 52-53) pateikti rezultatai logit ir cloglog modelių atskirai vyrams ir moterims. Lyginant pagal modelio tikimo duomenims statistikas, vėlgi šiek tiek geresnis atrodo modelis su cloglog sąryšio funkcija.

Analizuojant parametrų įverčius galima pastebėti, kad verslininkų pasitenkinimą darbu tiek vyrams, tiek moterims teigiamai veikia „pažangios“ firmos turėjimas: pareigų pasikeičiamumo laipsnis, klientų poreikius analizuojanti paslauga bei faktorius „firmos savybės“. „Laisvalaikio“ faktorius taip pat teigiamai veikia pasitenkinimą darbu, tačiau moterims jis daug reikšmingesnis negu vyrams. Darbo laiko faktorius neigiamai veikia vyrų pasitenkinimą ir teigiamai – moterų, o atsilikusių darbų tvarkymas laisvalaikiu moterims reikšmingai pasitenkinimo darbu nekeičia, tuo tarpu vyrams jį mažina. Didesnį pasitenkinimą darbu jaučia nevedę arba išsiskykę, neturintys mokyklą lankančių vaikų, turintys aukštąjį išsilavinimą, vyresni vyrai. Tuo tarpu moterys labiau patenkintos darbu yra ištekėjusios, jaunesnės negu 44 metų ir neturinčios į mokyklą einančių vaikų. Labiau patenkinti darbu yra patys pasirinkę dirbti sektoriuje bei siekę profesinės realizacijos asmenys. Pinigai, kaip priežastis tapti verslininku teigiamai veikia moterų ir neigiamai vyrų pasitenkinimą (lyginant su savarankišku gyvenimo organizavimu kaip priežastimi).

IŠVADOS

Atlikusi Italijos Veneto regiono verslininkų pasitenkinimo darbu daugiamatę analizę, naudodama kategoriniams duomenims tinkamus metodus galiu pateikti šias pagrindines darbo išvadas :

1. Ranginio priklausomojo kintamojo analizei galima rinktis šiuos metodus:
 - a) tiesinę regresiją su optimaliuoju skalės transformavimu, tačiau šis metodas dėl savo veikimo principo labiau tinka didesnių kategorijų skaičių turintiems kitamiesiems ir mūsų keturių rangų kintamajam jis netaikytinas;
 - b) logistinę regresiją, prieš tai priklausomąjį kintamąjį sudvireikšminus, tačiau bet kokios kintamųjų transformacijos yra informacijos praradimas ir taikytinos nesugebant išspręsti problemos kitokiais būdais;
 - c) ranginę regresiją, tačiau prieš tai būtina patikrinti lygiagrečiųjų tiesių prielaidą ir jei ši prielaida galioja, būtent toks būdas ranginio kintamojo analizei yra labiausiai rekomenduojamas;
 - d) daugiareikšmę (multinominę) logistinę regresiją, tačiau šis metodas rekomenduojamas ir prasmingas yra tik kai negalioja lygiagrečiųjų tiesių prielaida.

2. Atsiradus faktorinės analizės reikmei, kai turimi kategoriniai ranginiai arba dvireikšmiai kintamieji, Pearson'o korelacių matricą, reikalingą faktorių išskyrimui, galima bandyti pakeisti jos aproksimuotų reikšmių, t.y. polichorinių ir tetrachorinių korelacių matrica.

Analizuojant pasitenkinimą darbu, galima išskirti šiuos pagrindinius Italijos smulkiųjų verslininkų pasitenkinimo darbu veiksnius:

- 'pažangios' firmos turėjimas : pareigų pasikeičiamumo laipsnis, klientų poreikius analizuojanti paslauga bei faktorius 'firmos savybės' teigiamai veikia tiek vyrų, tiek moterų pasitenkinimą darbu;
- turiningo laisvalaikio turėjimas taip pat teigiamai susijęs su pasitenkinimu darbu, tačiau tai daug reikšmingesnis faktorius moterims negu vyrams;
- darbo laikas yra neigiamas faktorius vyrų pasitenkinimui darbu ir teigiamai veikia moterų pasitenkinimą, o ir atsilikusių darbų tvarkymas laisvalaikiu moterims reikšmingai pasitenkinimo darbu nekeičia, tuo tarpu vyrams jį mažina;
- didesnį pasitenkinimą darbu jaučia nevedę arba išsiskybę, neturintys mokyklą lankančių vaikų, vyresni vyrai su aukštesniu išsilavinimu, tuo tarpu moterys labiau patenkintos darbu yra ištekėjusios, jaunesnės negu 44 metų ir neturinčios į mokyklą einančių vaikų;
- verslo pradėjimo motyvai taip pat susiję su pasitenkinimu darbu: lyginant su savarankišku gyvenimo pradėjimu kaip bazine verslo pradėjimo priežastimi, profesinė realizacija yra teigiamas pasitenkinimo darbu veiksnys, o pinigai teigiamai veikia moterų ir neigiamai – vyrų pasitenkinimą darbu.

LITERATŪRA

- C. Agostinelli e S. Sartorelli, *Introduzione al linguaggio di SAS System*, Quaderni ASID, Facolta di Scienze Statistiche, Universita degli Studi di Padova (2002)
- Baltatescu Sergiu, *Problems of transforming scales of life satisfaction*, Euromodule workshop Berlin (2002)
- Borzaga Carlo, *Qualita del lavoro e soddisfazione dei lavoratori nei servizi sociali: un'analisi comparata tra modelli di gestione*, Working Paper n.9 (2000)
- Cano Jamie, Miller Greg, *A Gender Analysis of Job Satisfaction, Job Satisfier Factors, and Job Dissatisfier Factors of Agricultural Education Teachers*, Journal of Agricultural Education (2001)
- Čekanavičius V., Murauskas G. , *Statistika ir jos taikymai II*, Vilnius: TEV (2004).
- Chau-Kuang Chen, John Hughes Jr., *Using Ordinal Regression Model to Analyze Student Satisfaction Questionnaires*, Paper presented at the Annual Forum of the Association for Institutional Research (2004)
- Fabbris Luigi, *Appunti della Statistica Sociale(I)*, Dipartimento di Scienze Statistiche, Universita di Padova (2006)
- Friendly Michael, *Categorical Data Analysis with Graphics*, SCS Short Course, York University (2007)
- Giudici Paolo , *Data mining: metodi statistici per le applicazioni aziendali* , Milano: McGraw-Hill (2001)
- J. Juras, Z. Pasarić, *Application of tetrachoric and polychoric correlation coefficients to forecast verification*, Geofizika , Vol. 23 No. 1 (2006)
- G. Pierracini, *Gestione e trattamento dei dati con il SAS*, Dispensa disponibile alla pagina del corso (2004)
- G. Pierracini, *Programmazione avanzata e analisi dei dati con il SAS*, Dispensa disponibile alla pagina del corso (2004)
- SAS Institute Inc., *SAS Procedures Guide*, Cary, NC, USA (2003)
- SAS Institute Inc., *SAS STAT User's Guide*, Cary, NC, USA (2003)
- SAS Textbook Examples, *An Introduction to Categorical Analysis by Alan Agresti*, (2004)
- Scott. M. Lynch, *Multicollinearity, Soc 504* (2003)
- SPSS, Inc., *Ordinal Regression Analysis*, SPSS Advanced Models 10.00., Chicago, IL (2002)
- Verbeek Marno , *Econometria*, Bologna: Zanichelli (2006)
- Vittadini Giorgio , *Valutazioni in sanita e quantificazione delle variabili*, Milano (2005)

PRIEDAI

1. KINTAMIEJI IR JŲ APRAŠOMOSIOS STATISTIKOS

Priklausomasis kintamasis : PASITENKINIMAS DARBU

Kategorijos	Dažnis			Procentinis dažnis		
	Bendras	Vyrai	Moterys	Bendras	Vyrai	Moterys
Nepatenkinti	79	23	56	6,5	5,9	6,8
Nei patenkinti, nei nepatenkinti	120	36	84	9,9	9,2	10,2
Pakankamai patenkinti	438	146	292	36,0	37,4	35,4
Patenkinti	578	185	393	47,6	47,4	47,6
IŠ VISO	1215	390	825	100,0	100,0	100,0

Kiekybiniai kintamieji

Kintamasis	Min	Max	Vidurkis	St. nuokrypis
Amžius	21	83	48,14	10,80
Šeimos narių skaičius	0	11	3,40	1,09
Vaikų skaičius	0	6	1,51	1,07
Vaikų, gyvenančių kartu skaičius	0	5	1,17	1,00
Val. skaičius per d. darbui	0	18	9,51	2,47
Val. skaičius per d. sau/hobiams	0	7	,8	0,95
Valandų sk. per d. poilsiui	0	15	7,4	1,2
Atostogų d. sk. per praeitus metus	0	120	13,77	11,31
Vidut. dienų komandiruotėje sk. per mėn.	0	20	0,48	1,80
Darbuotojų skaičius	0	184	7,52	11,86
Lengvatų darbuotojams skaičius *	0	2	0,13	0,38

* Apskaičiuotas kaip suma šių paslaugų firmos darbuotojams: valgykla arba dienpinigiai maistui, vaikų darželis, transporto paslaugos, stipendijos darbuotojų vaikų studijoms

Kokybiniai kintamieji (1)

Kintamasis	Kategorija	Dažniai (%)		
		Bendras	Vyrai	Moterys
LYTIS	Moteris	67,9		
	Vyras	32,1		
ŠEIMYBINĖ PADĖTIS	Nesituokęs/usi	14,7	13,9	15,1
	Susituokęs/usi	74,9	81,8	71,7
	Išsiskyęs/usi	6,2	3,6	7,4
	Našlys/ė	4,2	0,8	5,8
TURI MOKYKLINIO AMŽIAUS VAIKŲ	Taip	41,5	38,7	42,9
	Ne	58,5	61,4	57,1
IŠSILAVINIMAS	Pradinis	12,1	14,9	10,8
	Vidurinis	36,3	40,0	34,6
	Profes. mokykla	12,1	14,4	11,0
	Aukštesnysis	30,9	24,9	33,8
	Aukštasis	8,5	5,9	9,7
LANKĖ TOBUL. KURSUS	Taip	17,5	13,6	19,4
	Ne	82,5	86,4	80,6
INVESTUOJA Į PERSONALO TOBUL.	Taip	6,0	6,2	5,9
	Ne	94,0	93,8	94,1
INVESTUOJA Į FIRMOS INFORM. SIST. PLĖTRĄ	Taip	6,0	5,6	6,4
	Ne	94,0	94,4	93,6
PERSONALO PAREIGŲ PASIKEIČIAMUMO LAIPSNIS FIRMOJE	Aukštas	20,4	25,4	18,2
	Gana aukštas	49,1	45,1	51,0
	Nėra	30,4	29,5	30,8
KOKYBĖS KONTROLĖS PATVIRTINIMO BŪKLĖ	Patvirtinta	13,0	15,6	11,7
	Ne	6,4	5,9	6,7
	Nesvarbi firmos tikslams	80,6	78,5	81,6
TURI KLIENTŲ POREIKIŲ ANALIZĖS PASLAUGĄ	Taip	6,4	5,9	6,7
	Ne	93,6	94,1	93,3
FIRMA TURI TINKLALAPĮ	Taip	32,7	25,9	36,0
	Ne	67,3	74,1	64,0
PREKIŲ APYVARTOS POKYČIAI LYGINANT SU PRAĖJUSIAIS METAIS	Padidėjo	28,6	32,3	26,9
	Nepasikeitė	35,8	35,1	36,1
	Sumažėjo	35,6	32,6	37,1
DARBO SEKTORIUJE PRIEŽASTYS	Pasirinkimas	72,3	79,0	69,1
	Reikalingumas	22,4	15,4	25,8
	Prievolė	5,3	5,6	5,1

Kokybiniai kintamieji (2)

Kintamasis	Kategorija	Dažniai (%)		
		Bendras	Vyrai	Moterys
TAPIMO VERSLININKU/E MOTYVAI	Pajamos	9,7	10,0	9,5
	Socialinis prestižas	2,9	4,9	2,0
	Profesinė realizacija	23,0	26,0	21,7
	Savarankiškas gyven. organizavimas	64,4	59,1	66,9
FIRMA ĮKŪRĖ PATS/PATI	Taip	53,3	59,7	50,2
	Ne	46,7	40,3	49,8
FIRMAI ĮKURTI PANAUDOTAS NUOSAVAS KAPITALAS	Taip	26,3	24,4	27,1
	Ne	73,7	75,6	72,9
TURI LAISVO NUO DARBO LAIKO	Taip	49,3	49,0	48,4
	Ne	50,7	50,0	51,6
SPORTUOJA	Taip	49,8	56,8	46,3
	Ne	50,2	43,2	53,7
LANKO DRAUGUS, GIMINAIČIUS	Taip	88,4	88,9	88,2
	Ne	11,6	11,1	11,8
TENKA KELIAUTI DARBO REIKALAIS	Taip	12,3	12,3	12,4
	Ne	87,7	87,7	87,6
DIRBA IR SAVAITGALIAIS	Taip, ir šešt, ir sekm.	18,0	15,6	19,1
	Taip, šešt. arba sekm.	44,1	55,6	38,6
	Ne	37,9	28,7	42,3
LAISVU LAIKU TVARKO ATSILIKUSIUS DARBO REIKALUS	Taip	61,1	57,4	62,9
	Ne	38,9	42,6	37,1
KAIP ĮSIVAIZDUOJA PADĖTĮ SEKTORIUJE ATEITYJE	Pagerės	22,7	22,3	22,9
	Nepasikeis	33,8	30,0	35,6
	Pablogės	43,5	47,7	41,5

Transformuotieji kintamieji

Kintamasis	Kategorija	Dažniai (%)		
		Bendras	Vyrai	Moterys
ŠEIMYNINĖ PADĖTIS	Nesusituokęs/usi	14,7	13,9	15,1
	Susituokęs/usi	74,9	81,8	71,7
	Išsiskykęs/usi arba našlys/ė	10,8	4,4	13,2
DARBO SEKTORIUJE PRIEŽASTYS	Pasirinkimas	72,3	79,0	69,1
	Reikmė arba prievolė	27,7	21,0	30,9
PREKIŲ APYVARTOS POKYČIAI LYGINANT SU PRAĖJUSIAIS METAIS	Padidėjo arba nepasikeitė	64,4	67,4	62,9
	Sumažėjo	35,6	32,6	37,1
AMZIAUS GRUPĖS	<35	11,1	9,2	12,0
	35-44	26,4	24,1	27,5
	45-54	33,9	30,3	35,6
	55-64	22,0	25,9	20,2
	>64	6,6	10,5	4,7
DARBUOTOJŲ SKAIČIUS GRUPĖM	<5	55,26	58,0	54,0
	5-10	27,8	26,9	28,2
	11-20	11,6	10,5	12,1
	>20	5,35	4,6	5,7
DARBUOTOJAI TURI BENT VIENĄ LENGVATĄ	Taip	15,35	19,2	13,6
	Ne	84,65	80,8	86,4
DARBO VALANDŲ SKAIČIUS	<8	12,8	4,6	16,6
	8-9	38,7	34,1	40,8
	10-12	39,9	51,8	34,3
	>12	8,7	9,5	8,4
VALANDŲ SKAIČIUS SAU, HOBIAMS	0	44,8	41,0	46,5
	1	34,2	36,0	33,3
	>1	21,1	23,0	20,2
ATOSTOGŲ SKAIČIUS PER METUS	0	17,0	17,2	17,0
	1-7	15,9	17,2	15,3
	8-14	16,7	19,0	15,6
	15-21	34,2	32,3	35,1
	>21	16,2	14,4	17,1

2. MULTIKOLINEARUMO TYRIMAS

Dalinės koreliacijos , VIF ir tolerancija

KINTAMASIS	DF	I Tipo dalinė kv. koreliacija	II Tipo dalinė kv. koreliacija	Tolerancija	VIF
1. Lytis	1	0.03345	0.06523	0.70759	1.41324
2. Tobulinimosi kursai	1	0.00041635	0.00140	0.86135	1.16097
3. Issilavinimas	1	0.00004138	0.00897	0.35027	2.85496
4. Seimyninė padėtis	1	0.00071367	0.00050295	0.77459	1.29101
5. Seimos nariu skaičius	1	0.11806	0.00464	0.08376	11.93831
6. Vaiku skaičius	1	0.01182	0.01069	0.11085	9.02107
7. Vaiku, gyvenančių kartu skaičius	1	0.01758	0.06915	0.09421	10.61513
8. Yra mokykla lankančių vaiku	1	0.00127	0.00119	0.17046	5.86636
9. Tapimo verslininku/e motyvai	1	0.01439	0.00011630	0.47093	2.12345
10. Darbo sektoriuje priežastys	1	0.03153	0.09190	0.49870	2.00523
11. Firma ikure pats/pati	1	0.00081533	0.01244	0.23684	4.22211
12. Firmai ikurti panaudotas nuosavas kapitalas	1	0.00114	0.02133	0.17194	5.81586
13. Personalo pareigų pasikeičiamumo laipsnis	1	0.01290	0.00053711	0.28811	3.47090
14. Klientų poreikių analizės paslaugos buvimas	1	0.00000328	0.00092134	0.38572	2.59255
15. Investavimo į firmos informacinę sistemą bukle	1	0.04246	0.00931	0.23747	4.21111
16. Investavimo į personalo tobulinimą bukle	1	0.01810	0.00880	0.19573	5.10907
17. Produktų kokybės sertifikavimo busena	1	0.18972	0.01550	0.30378	3.29190
18. Firmos tinklalapio egzistavimas	1	0.02768	0.22595	0.38783	2.57847
19. Darbuotojų skaičius	1	0.02696	0.00019451	0.37167	2.69059
20. Lengvatų darbuotojams skaičius	1	0.00842	0.07733	0.38448	2.60092
21. Tenka keliauti darbo reikalais	1	0.02668	0.00162	0.29153	3.43012
22. Dienų skaičius komandiruotėje	1	0.00625	0.00254	0.19910	5.02272
23. Kaip išsivaizduoja ateities situacija sektoriuje	1	0.02666	0.00450	0.34873	2.86755
24. Amžius	1	0.09708	0.08786	0.33156	3.01600
25. Prekių apyvartos pokyčiai lyginant su praeitais metais	1	0.10707	0.23404	0.44736	2.23533
26. Darbo valandų per dieną skaičius	1	0.09630	0.13782	0.41018	2.43794
27. Dirba ir savaitgaliais	1	0.00457	0.01004	0.34039	2.93784
28. Laisvų laikų tvarko atsiliusius darbo reikalus	1	0.01584	0.03873	0.26108	3.83027
29. Turi laisvo nuo darbo laiko	1	0.01175	0.00933	0.27796	3.59768
30. Valandų sau, hobiams skaičius per dieną	1	0.07169	0.10548	0.34405	2.90657
31. Valandos poilsiu	1	0.00866	0.00933	0.04830	20.70295
32. Atostogų dienos praeitais m.	1	0.02610	0.03404	0.33563	2.97948
33. Sportuoja	1	0.00177	0.02244	0.27997	3.57178
34. Lanko draugus, giminačius	1	0.06866	0.04865	0.54435	1.83704

STIPRIAUS SUSIJUSIŲ KINTAMŲJŲ BEIEŠKANT:

SPEARMAN‘O KORELIACIJOS

1-OJI SUSIJUSIŲ KINTAMŲJŲ GRUPĖ

KINTAMASIS	Seimos nariu skaicius	Vaiku skaicius	Vaiku, gyvenanciu kartu skaicius	Yra mokykla lankanciu vaiku	Amziaus grupes	Issilavinimas	Seimynine padetis
Seimos nariu skaicius	1	0.5431499778	0.7274136218	0.3977802551	0.0548017744	0.0399926496	0.095542753
Vaiku skaicius	0.5431499778	1	0.6890837941	0.5097397811	-0.422336389	0.2479103985	-0.352292319
Vaiku, gyvenanciu kartu skaicius	0.7274136218	0.6890837941	1	0.4637915686	-0.005389035	0.0630662621	-0.300578826
Yra mokykla lankanciu vaiku	0.3977802551	0.5097397811	0.4637915686	1	0.1830963726	-0.079808413	-0.277907365
Amziaus grupes	0.0548017744	-0.422336389	-0.005389035	0.1830963726	1	-0.285665635	0.3102229104
Issilavinimas	0.0399926496	0.2479103985	0.0630662621	-0.079808413	-0.285665635	1	-0.136507101
Seimynine padetis	0.095542753	-0.352292319	-0.300578826	-0.277907365	0.3102229104	-0.136507101	1

2-OJI SUSIJUSIŲ KINTAMŲJŲ GRUPĖ

KINTAMASIS	Firma ikure pats/pati	Firmai ikurti panaudotas nuosavas kapitalas	Darbo valandu per diena skaicius	Valandu sau,hobiams skaicius per diena	Turi laisvo nuo darbo laiko	Lanko draugus, gimnaisius	Amziaus grupes
Firma ikure pats/pati	1	0.6392232717	0.1199101755	-0.107731665	-0.053424041	-0.045341993	-0.180847574
Firmai ikurti panaud. savas kapitalas	0.6392232717	1	0.0285253869	0.1228810435	0.1731038898	0.1515566635	-0.037058065
Darbo val. per diena skaicius	0.1199101755	0.0285253869	1	-0.302617502	-0.191671138	-0.18211819	-0.011054118
Valandu sau, hobiams sk. per diena	-0.107731665	0.1228810435	-0.302617502	1	0.3960747867	0.389100328	0.0821577274
Turi laisvo nuo darbo laiko	-0.053424041	0.1731038898	-0.191671138	0.3960747867	1	0.2915217663	0.0845826163
Lanko draugus, gimnaisius	-0.045341993	0.1515566635	-0.18211819	0.389100328	0.2915217663	1	0.0877270645
Amziaus grupes	-0.180847574	-0.037058065	-0.011054118	0.0821577274	0.0845826163	0.0877270645	1

SPEARMAN'O KORELIACIJOS

3-IOJI SUSIJUSIŲ KINTAMŲJŲ GRUPĖ

KINTAMASIS	Invest. i firmos informacine sistema	Invest. i personalo tobulinima	Kokybes kontroles patvirtinimas	Darbuotoju skaicius	Darbuotojai turi lengvatu	Firma turi savo tinklalapi
Invest. i firmos informacine sistema	1	0.7820459052	0.0676697772	0.0145070792	0.073674934	0.039721157
Invest. i personalo tobulinima	0.7820459052	1	0.0017360941	0.0042239544	0.036811787	0.059811691
Kokybes kontroles patvirtinimas	0.0676697772	0.0017360941	1	0.2512001814	0.192656398	0.235636916
Darbuotoju skaicius	0.0145070792	0.0042239544	0.2512001814	1	0.289044367	-0.323287522
Darbuotojai turi lengvatu	0.073674934	0.036811787	0.192656398	0.289044367	1	0.1716050367
Firma turi savo tinklalapi	0.039721157	-0.059811691	0.235636916	-0.323287522	0.1716050367	1

4-OJI SUSIJUSIŲ KINTAMŲJŲ GRUPĖ

KINTAMASIS	Tenka keliauti darbo reikalais	Dienu skaicius komandiruotej	Darbuotoju skaicius	Lengvatu sk. darbuotojams	Firma turi savo tinklalapi	Issilavinimas
Tenka keliauti darbo reikalais	1	0.7235845703	0.1713960343	0.1001296971	0.2126738898	-0.232810224
Dienu skaicius komandiruotej	0.7235845703	1	0.1268972831	0.1132852957	0.1590251427	-0.172606993
Darbuotoju skaicius	0.1713960343	0.1268972831	1	0.3635270132	0.2997826621	-0.162311135
Lengvatu sk. darbuotojams	0.1001296971	0.1132852957	0.3635270132	1	0.1775873888	-0.131098185
Firma turi savo tinklalapi	0.2126738898	0.1590251427	0.2997826621	0.1775873888	1	-0.270183657
Issilavinimas	-0.232810224	-0.172606993	-0.162311135	-0.131098185	-0.270183657	1

SPEARMAN'Ų KORELIACIJOS

5-OJI SUSIJUSIŲ KINTAMŲJŲ GRUPĖ

KINTAMASIS	Turi laisvo nuo darbo laiko	Valandos poilsiui	Laisvu laiku tvarko atsilikusius darbo reikalus	Valandos sau, hobiams	Sportuoja	Lanko draugus, giminaicius
Turi laisvo nuo darbo laiko	1	0.7384573327	-0.39493824	0.422515035	0.2395629397	0.1915875663
Valandos poilsiui	0.7384573327	1	-0.492945558	0.80924719	0.5790729397	0.4915217663
Laisvu laiku tvarko atsilikusius darbo reikalus	-0.39493824	-0.492945558	1	0.3627249847	-0.395310561	-0.328614108
Valandos sau, hobiams	0.422515035	0.80924719	-0.3627249847	1	0.316090219	0.217390983
Sportuoja	0.2395629397	0.5790729397	-0.395310561	0.316090219	1	0.288930244
Lanko draugus, giminaicius	0.1915875663	0.4915217663	-0.328614108	0.217390983	0.2388930244	1

KINTAMUJŲ TRANSFORMACIJOS

Kintamasis	Kategorija	Dažniai (%)		
		Bendras	Vyrai	Moterys
INVESTUOJA Į PERSONALO TOBUL.	Taip	6,0	6,2	5,9
	Ne	94,0	93,8	94,1
INVESTUOJA Į FIRMOS INFORM. SIST. PLĖTRĄ	Taip	6,0	5,6	6,4
	Ne	94,0	94,4	93,6



INVESTUOJA Į PERSONALO TOBULINIMĄ IR/ARBA INFORM. SIST. PLĖTRĄ	Investuoja į abu	3,1	3,1	3,2
	Investuoja arba į personalo tobulinimą, arba inform. sist. plėtrą	5,9	5,6	6,0
	Ne	91,0	91,3	90,8

3. RANGINĖ LOGISTINĖ REGRESIJA (link=LOGIT)

KINTAMASIS		DF	B	S. E.	P> Chi-square	Exp(B)
Intercept	3	1	0.6607	0.0329	<.0001	1.936
Intercept	2	1	2.9595	0.0342	<.0001	19.288
Intercept	1	1	4.1538	0.0358	<.0001	63.676
Moteris	0	1	0.2128	0.00828	<.0001	1.237
Sportuoja	1	1	-0.1696	0.0102	<.0001	0.844
Lanko draugus, gimines	1	1	-0.8143	0.0195	<.0001	0.443
ŠEIM. PADETIS <i>(bazinė=nesusituokęs/usi)</i>						
<i>Išsiskyres/usi arba našlys/ė</i>	1	1	-0.1843	0.0202	<.0001	0.832
<i>Susituokęs, -usi</i>	2	1	0.3343	0.0135	<.0001	1.397
Lankė tobulinimosi kursus	1	1	0.2533	0.00942	<.0001	1.288
INVEST. Į PERSONALĄ, INFORM. SIST. <i>(bazinė=jokių)</i>						
<i>Ir į personalą, ir į inform. sistemą</i>	0	1	0.1482	0.0182	<.0001	1.160
<i>Arba į personalą, arba į inform. sistemą</i>	1	1	-0.3934	0.0242	<.0001	0.675
Prekių apyvarta išaugo	1	1	0.0933	0.00728	<.0001	1.098
Darbas sektoriuje- pasirinkimas	1	1	0.4588	0.00802	<.0001	1.582
PAREIGŲ PASIKEIČIAMUMO LAIPSNIS FIRMOJE <i>(bazinis=nėra)</i>						
<i>Aukštas</i>	1	1	0.4781	0.0112	<.0001	1.613
<i>Gana aukštas</i>	2	1	-0.1869	0.00953	<.0001	0.829
Turi klientų reikmes analiz. paslaugą	1	1	0.3516	0.0187	<.0001	1.421
KOKYBĖS KONTROLĖS PATVIRTINIMAS <i>(bazinis= firmai nesvarbus)</i>						
<i>Neturi</i>	1	1	-0.0798	0.0179	<.0001	0.923
<i>Patvirtinta</i>	2	1	0.2031	0.0244	<.0001	1.225
Firmą įkūrė pats/pati	1	1	0.2397	0.00744	<.0001	1.271

KINTAMASIS		DF	B	S. E.	P> Chi-square	Exp(B)
Turi laisvo nuo darbo laiko	1	1	0.9237	0.0216	<.0001	2.519
Tenka keliauti darbo reikalais	1	1	-0.1620	0.0116	<.0001	0.850
KAIP ĮSIVAIZDUOJA ATEITIES SITUACIJĄ SEKTORIUJE (bazinė=pablogės)						
<i>Pagerės</i>	1	1	0.1998	0.0112	<.0001	1.221
<i>Nepasikeis</i>	2	1	0.2712	0.0103	<.0001	1.312
VERSLO PRADĖJIMO MOTYVAS (bazinė=savarankiškas gyvenimo organizavimas)						
<i>Pinigai</i>	1	1	-0.4120	0.0173	<.0001	0.662
<i>Socialinis prestižas</i>	2	1	0.0722	0.0265	0.0064	1.075
<i>Profesinė realizacija</i>	3	1	0.7187	0.0154	<.0001	2.052
Firma turi savo tinklalapį	1	1	0.1372	0.00908	<.0001	1.147
DIRBA IR SAVAITGALIAIS (bazinė=ne)						
<i>Taip, ir šeštadieniais ir sekmadieniais</i>	1	1	0.0900	0.0127	<.0001	1.094
<i>Taip, šeštadieniais arba sekmadieniais</i>	2	1	0.1510	0.00952	<.0001	1.163
IŠSILAVINIMAS (bazinis=pradinis)						
<i>Aukštasis</i>	1	1	0.3444	0.0252	<.0001	1.411
<i>Aukštesnysis</i>	2	1	-0.0297	0.0149	0.0463	0.971
<i>Profesinė mokykla</i>	3	1	0.00156	0.0173	0.9284	1.002
<i>Vidurinė mokykla</i>	4	1	-0.3281	0.0123	<.0001	0.720
Laisvu laiku tvarko atsilikusius darbo reikalus	1	1	-0.1433	0.0125	<.0001	0.866
DARBUOTOJŲ SKAIČIUS (bazinis= '<5')						
<i>>20</i>	1	1	0.9254	0.0386	<.0001	2.523
<i>11-20</i>	2	1	-0.4733	0.0230	<.0001	0.623
<i>5-10</i>	3	1	0.1090	0.0184	<.0001	1.115

KINTAMASIS		DF	B	S. E.	P> Chi-square	Exp(B)
ATOSTOGU DIENU SKAICIUS PER METUS (bazinis=0)						
>21	1	1	0.3789	0.0166	<.0001	1.461
15-21	2	1	0.0641	0.0119	<.0001	1.066
8-14	3	1	0.0273	0.0144	0.0576	1.028
1-7	4	1	-0.4362	0.0144	<.0001	0.646
Firmos darbuotojai turi lengvatų	1	1	0.1943	0.0109	<.0001	1.214
AMŽIUS (bazinis=, <35')						
>64	1	1	0.1214	0.0221	<.0001	1.129
55-64	2	1	-0.1455	0.0142	<.0001	0.865
45-54	3	1	-0.0854	0.0130	<.0001	0.918
35-44	4	1	-0.1494	0.0139	<.0001	0.861
DARBO VAL. PER DIENĄ (bazinis=, <8')						
>12	1	1	0.5805	0.0218	<.0001	1.787
10-12	2	1	-0.2230	0.0122	<.0001	0.800
8-9	3	1	-0.1576	0.0129	<.0001	0.854
VALANDŲ SKAIČIUS SAU (bazinis=0)	1	1	0.2415	0.0119	<.0001	1.273
>1						
1	2	1	-0.0103	0.0102	0.3140	0.990

4. FAKTORINĖ ANALIZĖ (1) - ALTERNATYVUS METODAS

POLICHORINIŲ- TETRACHORINIŲ KORELIACIJŲ MATRICA

KINTAMASIS	Tenka keliauti darbo reikalais	Sportuoja	Firma turi savo tinklalapi	Darbuotojai turi lengvatu	Kokybes kontroles busena	Dirba savaitgaliais	Darbuotoju skaicius	Darbo val skaicius	Valandu skaicius sau	Lanko draugus, giminaicius	Atostogu skaicius	Turi laisvo nuo darbo laiko
Tenka keliauti darbo reikalais	1	0.1239	0.2722	0.1342	0.0917	0.0688	0.2539	0.1038	0.0629	0.1264	0.0651	0.1506
Sportuoja	0.1239	1	0.0698	0.0447	0.0701	-0.0593	0.0186	-0.1241	0.3741	0.4388	0.1745	0.3791
Firma turi savo tinklalapi	0.2722	0.0698	1	0.2416	0.3347	-0.0213	0.4232	0.0481	0.0514	0.0657	0.1221	0.0839
Darbuot. turi lengvatu	0.1342	0.0447	0.2416	1	0.2726	0.00821	0.4548	0.0394	0.0495	0.0271	0.1025	0.0227
Kokybes kontroles busena	0.0917	0.0701	0.3347	0.2726	1	0.0056	0.3212	0.0193	0.0331	0.0498	0.0205	0.0544
Dirba savaitgaliais	0.06887	-0.0593	-0.0213	0.0082	0.0056	1	-0.0841	0.3036	-0.0962	-0.0397	-0.2027	-0.0474
Darbuotoju skaicius	0.2539	0.0186	0.4232	0.4548	0.3212	-0.0841	1	0.0792	-0.0189	-0.0173	0.0925	0.0196
Darbo val skaicius	0.1038	-0.1241	0.0481	0.0394	0.0193	0.3036	0.0792	1	-0.4855	-0.2761	-0.1627	-0.2859
Valandu skaicius sau	0.06298	0.3741	0.0514	0.0495	0.0331	-0.0962	-0.0189	-0.4855	1	0.4173	0.1433	0.4745
Lanko draugus, giminaicius	0.1264	0.4388	0.0657	0.0271	0.0498	-0.0397	-0.0173	-0.2761	0.4173	1	0.1348	0.5735
Atostogu skaicius	0.0651	0.1745	0.1221	0.1025	0.0205	-0.2027	0.0925	-0.1627	0.1433	0.1348	1	0.1531
Turi laisvo nuo darbo laiko	0.1506	0.3791	0.08394	0.0227	0.0544	-0.0474	0.0196	-0.2859	0.4745	0.5735	0.1531	1

DUOMENŲ TINKAMUMO FAKTORINEI ANALIZEI MATAI

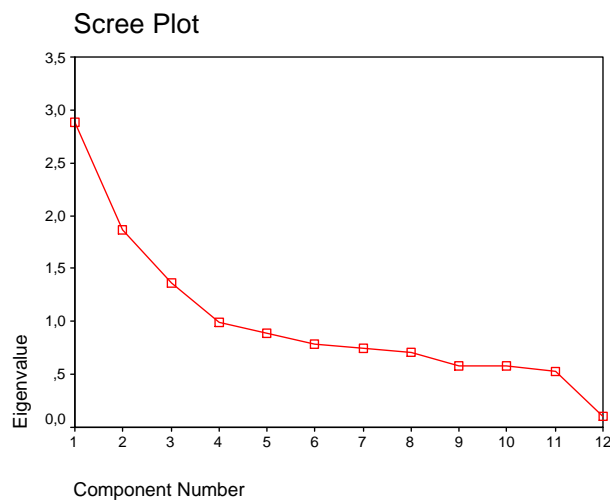
Atskiro kintamojo tikimas faktorinei analizei- MSA matai											
Turi laisvo laiko	Sportuoja	Lanko draugus	Tenka keliaut darbo reikalais	Turi tinklalapį	Suteikia lengvatų darbuotojams	Kokybės kontrolės patvirtinimas	Dirba savaitgaliais	Darbuotojų skaičius	Atostogų skaičius	Darbo valandų skaičius	Valandų sau/hobiams skaičius
0.6484	0.90987	0.65611	0.74582	0.71889	0.68876	0.73291	0.57868	0.64711	0.76875	0.66839	0.88517

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,708
Bartlett's Test of Sphericity	Approx. Chi-Square	3500,204
	df	66
	Sig.	,000

RENKAM FAKTORIŲ SKAIČIŲ

TIKRINĖS REIKŠMĖS IR PAAIŠKINTA DISPERSIJA				
	Tikrinė reikšmė	Skirtumas	Paaiškinta dispersija	Sukauptoji paaiškintoji dispersija
1	2.86244006	1.00210202	0.2385	0.2385
2	1.86033804	0.47340853	0.1550	0.3936
3	1.38692950	0.43062548	0.1156	0.5091
4	0.95630402	0.09163685	0.0797	0.5888
5	0.86466718	0.07957048	0.0721	0.6609
6	0.78509669	0.03533426	0.0654	0.7263
7	0.74976243	0.03494677	0.0625	0.7888
8	0.71481567	0.10693562	0.0596	0.8484
9	0.60788005	0.02324471	0.0507	0.8990
10	0.58463533	0.06107469	0.0487	0.9477
11	0.52356065	0.41999026	0.0436	0.9914
12	0.10357039		0.0086	1.0000



Faktorių svoriai

KINTAMIEJI	Factor1	Factor2	Factor3
Turi laisvo nuo darbo laiko	0.88840	-0.11475	0.23009
Sportuoja	0.72754	-0.06121	0.15994
Lanko draugus,giminačius	0.86414	-0.14350	0.23918
Tenka keliauti darbo reikalais	0.22826	0.40312	0.30384
Firma turi savo tinklalapį	0.19844	0.63600	-0.04969
Darbuotojai turi lengvatų	0.13720	0.54278	-0.14061
Kokybės kontrolės būseną	0.15991	0.54514	-0.05628
Dirba savaitgaliais	-0.17631	0.07501	0.73641
Darbuotojų skaičius	0.09725	0.72206	-0.16479
Atostogų skaičius	0.33210	0.11674	-0.44750
Darbo val skaičius	-0.34280	0.31537	0.59547
Valandu skaičius sau	0.62682	-0.15481	-0.09581

FAKTORIŲ PAAIŠKINTA DISPERSIJA		
Factor1	Factor2	Factor3
2.8624401	1.8603380	1.3869295

FAKTORIŲ SUKIMAS

ORTOGONALIOSIOS TRANSFORMACIJOS MATRICA			
	1	2	3
1	0.94249	0.20975	-0.26023
2	-0.16933	0.97091	0.16930
3	0.28817	-0.11550	0.95059

INTERPRETAVIMAS

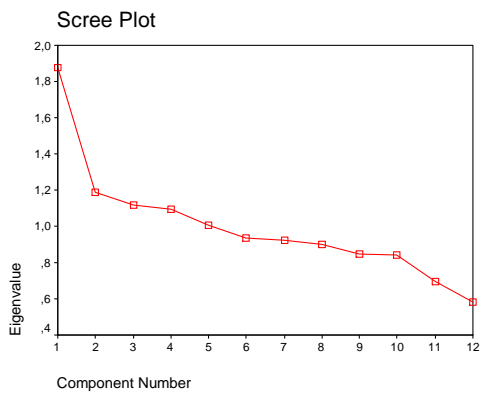
FAKTORIŲ SVORIAI			
	Factor1 (LAISVALAIKIS)	Factor2 (FIRMA)	Factor3 (DARBO LAIKAS)
Turi laisvo nuo darbo laiko	0.92305	0.04836	-0.03190
Sportuoja	0.74216	0.07470	-0.04765
Lanko draugus,giminaicius	0.90767	0.01431	-0.02180
Tenka keliauti darbo reikalais	0.23443	0.40418	0.29767
Firma turi savo tinklalapį	0.06501	0.66486	0.00880
Darbuotojai turi lengvatų	-0.00312	0.57201	-0.07747
Kokybės kontrolės būseną	0.04218	0.56932	-0.00282
Dirba savaitgaliais	0.03333	-0.04921	0.75860
Darbuotojų skaičius	-0.07810	0.74048	-0.05970
Atostogų skaičius	0.16428	0.23469	-0.49204
Darbo val skaičius	-0.20489	0.16552	0.70865
Valandu skaičius sau	0.58937	-0.00776	-0.28040

5. FAKTORINĖ ANALIZĖ (2) - TRADICINIS METODAS

Duomenų tikimo faktorinei analizei tikrinimas

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.620
Bartlett's Test of Sphericity	Approx. Chi-Square	586,110
	df	66
	Sig.	,000



	Tikrinė reikšmė	Paašikinta dispersija	Sukauptoji paašikinta dispersija
1	1.878	0.1563	0.1563
2	1.186	0.0989	0.2552
3	1.119	0.0933	0.3485
4	1,094	0.0911	0.4396
5	1,006	0.0838	0,5234
6	0.933	0.0778	0.6012
7	0.922	0.0768	0.6780
8	0.901	0.0751	0.7531
9	0.845	0.0705	0.8236
10	0.841	0.0701	0.8937
11	0.694	0.0578	0.9515
12	0.582	0.0485	1.0000

Faktorių svoriai

KINTAMIEJI	Factor1	Factor2	Factor3
Turi laisvo nuo darbo laiko	0.47843	-0.21875	0.43879
Sportuoja	0.42634	-0.16421	0.23418
Lanko draugus, giminaicius	0.39452	-0.13314	0.04782
Tenka keliauti darbo reikalais	0.14478	0.20412	0.22703
Firma turi savo tinklalapį	0.10084	0.54720	-0.14969
Darbuotojai turi lengvatų	0.13720	0.24998	-0.09561
Kokybės kontrolės būseną	0.19663	0.39782	-0.15647
Dirba savaitgaliais	-0.36682	0.15237	0.43691
Darbuotojų skaičius	0.05735	0.52023	-0.36356
Atostogų skaičius	0.23780	0.09751	-0.41088
Darbo val skaičius	-0.31270	0.30204	0.49783
Valandų skaičius sau	0.50653	-0.13921	-0.15432

Faktorių svoriai po sukimo

	Factor1	Factor2	Factor3
Turi laisvo nuo darbo laiko	0.52855	0.02336	-0.34782
Sportuoja	0.64244	0.01270	-0.14765
Lanko draugus, giminaicius	0.51407	0.03831	-0.00180
Tenka keliauti darbo reikalais	0.10713	0.10418	0.29537
Firma turi savo tinklalapį	0.04151	0.67596	0.01253
Darbuotojai turi lengvatų	-0.00513	0.31201	-0.07747
Kokybės kontrolės būseną	0.04218	0.41782	-0.01841
Dirba savaitgaliais	0.16353	-0.14526	0.65930
Darbuotojų skaičius	-0.01364	0.64928	-0.12703
Atostogų skaičius	0.16428	0.00476	-0.44935
Darbo val skaičius	-0.21152	0.16323	0.50865
Valandų skaičius sau	0.55413	-0.15820	-0.19047

6. MODELIO PAIEŠKA:

Ordinal logit, Ordinal cloglog, Binary logit, Regresija su Optimal scaling

KINTAMASIS		DF	Ranginė logit	Ranginė cloglog	Binary logit	Regresija su optimal scaling
Intercept	3	1	0.1804***	-0.4117***		
Intercept	2	1	2.4456***	0.9209***		
Intercept	1	1	3.7401***	1.5132***	2.2881***	0.0141
Factor1 (Laisvalaikis)		1	0.2593***	0.0425***	0.2908***	0.1115***
Factor2 (Firma)		1	0.3463***	0.1332***	0.2254***	0.1654***
Factor3 (Darbo laikas)		1	0.1575*	-0.0529***	-0.1864***	0.0501*
Lytis=moteris	0	1	0.1684***	0.0578**	0.1150***	0.0040
SEIMYN. PDT. (Bazine=Nesusituokes) Issiskyres/naslys	1	1	0.1694***	0.0631***	-0.1103***	0.0161
Susituokes	2	1	0.1895***	0.0723**	0.2877***	
Lanke kursus	1	1	0.0652***	0.0735**	-0.0405	-0.0723*
Yra mokykl. lank.vaiku	1	1	-0.0187	-0.0233*	-0.00824	-0.0231
INVEST. I PERSONALA AR INF. SIST (bazinis=jokiu) Inv. i abi	0	1	0.3157***	0.2052***	0.3352***	0.1343***
Investuoja arba I personala, arba inf. sist	1	1	-0.8688***	-0.6029***	-1.3662***	
Prekiu apyvarta isaugo/nepasikeite	1	1	0.1255***	0.1582***	0.2587***	0.0167
Dirbti sektoriuj pasirinko pats	1	1	0.5730***	0.2794***	0.7442***	0.1473**
PAREIGU PASIKEIC. LAIPSNIS FIRMOJ (bazinis=nera) Aukstas	1	1	0.5376***	0.2394***	0.2836***	0.0965***
Vidutinis	2	1	-0.1912***	-0.0749***	0.0534***	
Turi klientu poreikius analiz. paslauga	1	1	0.6065***	0.2672***	0.7982***	0.0747

KINTAMASIS		DF	Ranginė logit	Ranginė cloglog	Binary logit	Regresija su optimal scaling
Firma ikure pats	1	1	0.1371***	0.00757	-0.0517*	0.0032
MANO,KAD SITUACIJA SEKTORIUIJ (bazinis=pabloges) Pageres	1	1	0.1403***	0.1362***	0.2999***	0.0864**
Nepasikeis	2	1	0.3102***	0.1673***	0.4038***	
TAPIMO VERSL. MOTYVAI (bazinis=savar. gyven. org.) Pinigai	1	1	-0.2977***	-0.1496***	-0.6742***	-0.142***
Soc. prestizas	2	1	-0.0628	0.0128*	-0.2414***	
Profes. Realizac.	3	1	0.6811***	0.3262***	0.8513***	
ISSILAVINIMAS (bazinis=pradinis) Aukstasis	1	1	0.2528***	-0.00792	0.8053***	-0.0325
Aukstesnysis	2	1	0.0497*	0.0424**	0.3343***	
Profesine mokykla	3	1	0.0403	0.0466*	-0.3552***	
Vidurinis	4	1	-0.3216***	-0.1461***	-0.3816***	
Laisvu laiku tvarko atsilikusius darbo reikalus	1	1	-0.3362***	-0.1169***	-0.0508***	-0.0042
AMZIUS (bazinis=<35) >64	1	1	-0.0671*	0.1822***	0.4459***	0.0330
55-64	2	1	-0.1171***	-0.1107***	0.1621***	
45-54	3	1	-0.2899***	-0.1880***	-0.7860***	
35-44	4	1	0.0321	-0.0198	0.1188***	

Reikšmingumo lygmenys *** 0.001 ** 0.01 * 0.05

PRIELAUDŲ TIKRINIMAS IR TIKIMO DUOMENIMS STATISTIKOS

RANGINĖ REGRESIJA : Proporcionaliųjų tiesių prielaida (Parallel lines)

Ordinal Logit

Lygiagrečių tiesių hipotezė		
Chi-kvadrato	DF	Pr > ChiQuadr
70.3679	60	0.1693

Ordinal cloglog

Lygiagrečių tiesių hipotezė		
Chi-kvadrato	DF	Pr > ChiQuadr
67.3803	60	0.2394

Modelių tikimo duomenims statistikos

Koeficientas	Binary logit	Ordinal logit	Ordinal cloglog	Regresija su optimal scaling
<i>Cox ir Snell R²</i>	0.1317	0.2143	0.2487	
<i>Nagelkerke R²</i>	0.2219	0.2431	0.2752	
<i>AIC (tik intercept)</i>	856.628	2124.17	2124.17	
<i>AIC (int ir kint.)</i>	782.576	1988.12	1984.22	
<i>R²</i>				0.1713

Prognozių tikslumas

Binary Logit

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.6	Somers' D	0.416
Percent Discordant	29.0	Gamma	0.417
Percent Tied	0.4	Tau-a	0.262
Pairs	124978	c	0.708

Ordinal logit

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	74.0	Somers' D	0.484
Percent Discordant	25.6	Gamma	0.486
Percent Tied	0.4	Tau-a	0.135
Pairs	279617	c	0.742

Ordinal cloglog

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	79.0	Somers' D	0.582
Percent Discordant	20.7	Gamma	0.584
Percent Tied	0.3	Tau-a	0.162
Pairs	279617	c	0.791

RANGINĖ REGRESIJA VYRAMS IR MOTERIMS ATSKIRAI :

Logit ir Cloglog sąryšio funkcijos

KINTAMASIS		DF	Logit vyrams	Logit moterims	Cloglog vyrams	Cloglog moterims
Intercept	3	1	0.2696**	0.7024**	-0.1699**	-0.0584
Intercept	2	1	2.6479**	2.9444**	1.2585**	1.2957**
Intercept	1	1	3.9482**	4.3275**	1.8604**	1.9680**
Factor1 (Laisvalaikis)		1	-0.0126	0.2919**	0.0433**	0.1434**
Factor2 (Firma)		1	0.4347**	0.2301**	0.1633**	0.0813**
Factor3 (Darbo laikas)		1	-0.0733**	0.3246**	-0.0876**	0.1303**
SEIMYN. PDT. (Bazine=Nesusituokes) Issiskyres/naslys	1	1	0.3426**	-0.0362	0.3134**	-0.0642*
Susituokes	2	1	0.1269**	0.2872**	-0.0529**	0.1704**
Lanke kursas	1	1	-0.0235*	0.2881**	0.00306	0.2009**
Yra mokykl. lank.vaiku	1	1	0.0381**	-0.0781**	-0.0233**	-0.0431**
INVEST. I PERSONALA AR INF. SIST (bazinis=jokiu) Investuoja i abi	0	1	0.4812**	-0.6773**	0.2684**	-0.2611**
Investuoja arba į personala, arba į inf. sistemą	1	1	-1.0525**	-0.4758**	-0.7228**	-0.4398**
Prekiu apyvarta isaugo/nepasikeite	1	1	0.0821**	0.3496**	0.1226**	0.2023**
Dirbti sektoriuj pasirinko pats	1	1	0.5384**	0.6058**	0.2537**	0.3385**
PAREIGU PASIKEIC. LAIPSNIS FIRMOJ (bazinis=nera) Aukstas	1	1	0.6861**	0.1372**	0.3801**	0.0346*
Vidutinis	2	1	-0.2634**	0.0426*	-0.1347**	0.0975**
Turi klientu poreikius analiz. paslauga	1	1	0.9671**	0.1519**	0.5789**	0.0871**
Firma ikure pats	1	1	0.1635**	-0.0244	0.0489**	-0.0678**

KINTAMASIS		DF	Logit vyrams	Logit moterims	Cloglog vyrams	Cloglog moterims
MANO, KAD SITUACIJA SEKTORIUJ (bazinis=pabloges) Pageres	1	1	0.2448**	-0.0169	0.2058**	0.1229**
Nepasikeis	2	1	0.3042**	0.2858**	0.1414**	0.1621**
TAPIMO VERSL. MOTYVAI (bazinis=savar. gyven. org.) Pinigai	1	1	-0.4727**	0.1635**	-0.2378**	0.1371**
Soc. prestizas	2	1	0.00799	-0.3588**	0.00885	-0.1939**
Profes. Realizac.	3	1	0.7243**	0.5739**	0.3517**	0.3047**
ISSILAVINIMAS (bazinis=pradinis) Aukstasis	1	1	0.5207**	-0.3216**	0.2923**	-0.1915**
Aukstesnysis	2	1	0.0949**	-0.3426**	-0.0269*	-0.1983**
Profesine mokykla	3	1	-0.1312**	0.5464**	-0.0854**	0.3876**
Vidurinis	4	1	-0.3050**	-0.3691**	-0.1207**	-0.3078**
Laisvu laiku tvarko atsilikusius darbo reikalus	1	1	-0.4008**	-0.0168	-0.1668**	-0.0312*
AMZIUS (bazinis=<35) >64	1	1	0.00137	-0.3029**	0.2345**	-0.1458**
55-64	2	1	0.0994**	-0.5191**	0.0418**	-0.3273**
45-54	3	1	-0.2656**	-0.2558**	-0.1775**	-0.1802**
35-44	4	1	0.0263	0.0936*	0.0131	0.0334

Proporcionaliųjų tiesių prielaidos tikrinimas

<i>Link f-ja</i>	MOTERYS		VYRAI	
	logit	cloglog	logit	cloglog
<i>Chi-kv.(58 ll)</i>	70.8222	65.2566	76.3740	66.3309
<i>P>Chi-kv.</i>	0.1203	0.2392	0.0533	0.2117

Tikimo duomenims statistikos:

<i>Link f-ja</i>	MOTERYS		VYRAI	
	logit	cloglog	logit	cloglog
<i>Cox ir Snell R²</i>	0.1941	0.2122	0.2338	0.2564
<i>Nagelkerke R²</i>	0.2168	0.2370	0.2622	0.2876
<i>AIC (tik intercept)</i>	1434.983	1434.983	694.917	694.917
<i>AIC (int ir kint.)</i>	1356.349	1342.026	670.355	661.060

Prognozių tikslumas

Cloglog moterims

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.7	Somers' D	0.418
Percent Discordant	28.9	Gamma	0.420
Percent Tied	0.4	Tau-a	0.264
Pairs	126414	c	0.709

Logit moterims

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.0	Somers' D	0.403
Percent Discordant	29.6	Gamma	0.405
Percent Tied	0.4	Tau-a	0.255
Pairs	126414	c	0.702

Prognozių tikslumas

Logit vyrams

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	73.8	Somers' D	0.479
Percent Discordant	25.8	Gamma	0.481
Percent Tied	0.4	Tau-a	0.300
Pairs	30008	c	0.740

Cloglog vyrams

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	75.4	Somers' D	0.511
Percent Discordant	24.2	Gamma	0.513
Percent Tied	0.4	Tau-a	0.320
Pairs	30008	c	0.756