

TARTU ÜLIKOOL
MATEMAATIKA-INFORMAATIKATEADUSKOND
Matemaatika instituut
Matemaatika eriala

Vootele Rõtov
**Valikvastustega küsimustike
reliaabluse hindamine**
Bakalaureusetöö (6 EAP)

Juhendaja: Margus Niitsoo

TARTU 2014

Valikvastustega küsimustike reliaabluse hindamine

Vootele Rõtov

Lühikokkuvõte

Käesolevas bakalaureusetöös vaadeldakse erinevaid võimalusi valikvastustega küsimustike reliaabluse hindamiseks küsimustiku ühekordse läbiviimise põhjal.

Töö käigus antakse lugejale vajaminevad taustteadmised valikvastustega küsimustike ning nende karakteristikute kohta, defineeritakse reliaablus klassikalise testiteooria raamistikus ja antakse vajalik matemaatiline baas reliaabluse hinnangu-te tuletamiseks. Seejärel tuletatakse viis reliaabluse hinnangut, hinnatakse nende kasutatavust praktikas ja tutvustatakse autori loodud programme, mis sisaldab programme tuletatud hinnangute leidmiseks küsimustiku läbiviimisel saadud andmete põhjal. Töö viimases osas tutvustab autor üht võimalikku viisi küsimustike arvulise tõlgenduse parandamiseks.

Töö annab eesti keelse ülevaate klassikalise testiteooria alustest matemaatilisest vaatepunktist lähtudes, ning tutvustab praktikule erinevaid võimalusi testi reliaabluse hindamiseks, sealjuures tutvustatakse meetodeid, millest osad annavad täpsemaid tulemusi, kui hetkel enamlevinud meetodid.

VÕTMESÕNAD: psühhomeetria, klassikaline testiteooria, valikvastustega küsimustikud, reliaablus, reliaabluse hindamine, Cronbachi alfa, Guttmani lambdad, glb.

On the Reliability of Likert Scale Questionnaires

Vootele Rõtov

Abstract

This bachelor's thesis gives an overview of different possibilities of estimating the reliability of Likert scale questionnaires based on a single admission.

The thesis contains a short overview of the necessary background information about Likert scales and their characteristics, definition of reliability in the context of the Classical Test Theory and mathematical basis for deriving estimations of reliability. Also, five estimates are derived and their usability in practise is discussed. Afterwards, the author gives a short overview of a programming library that implements methods for finding derived estimates based on data gathered with the questionnaire. Finally, the author proposes an approach for improving the quantitative interpretation of a questionnaire.

The thesis gives an overview of the foundations of the Classical Test Theory in Estonian from a mathematical viewpoint and introduces different possibilities for estimating the reliability of a questionnaire, some of which give more exact estimates than the ones in widespread use today.

KEYWORDS: psychometrics, classical test theory, Likert scale questionnaires, reliability, estimating reliability, Cronbach's alfa, Guttman's lambdas, glb.

Sisukord

Sisukord	3
Sissejuhatus	5
1 Taustinfo ja probleemipüstitus	6
1.1 Küsimustik, küsimus ja Likerti skaala	6
1.2 Likerti skaala tõlgendamine intervallskaalana	6
1.3 Küsimustiku valiidsus	7
1.4 Küsimustiku reliaablus	8
1.5 Valiidsuse ja reliaabluse suhe	9
1.6 Sisemine järjepidevus	10
1.7 Probleemi püstitus	11
2 Reliaablus klassikalises testiteoorias	12
2.1 Reliaabluse definitsioon	12
2.2 Vajalikud taustteadmised reliaabluse hindamiseks	19
3 Reliaabluse alumised tőkked	24
3.1 Alumine tõke λ_1	24
3.2 Alumine tõke λ_2	25
3.3 Alumine tõke λ_3	29
3.4 Alumine tõke λ_4	31
3.5 Alumine tõke glb	32
3.6 Hinnangute kasutatavus praktikas	32
4 Programmiteek	33
5 Intervallskaal alternatiivne tõlgendamine	34
5.1 Vastuste tõlgendamine intervallskaalal	35
5.2 Reliaabluse hindamine	35
5.3 Lähenedmise korrektsus	36
5.4 Alternatiivse tõlgenduse matemaatiline sõnastus	37
5.5 Hinnang saadud tõlgendusele	41
Kokkuvõte	42
Viited	43
Lisad	45
Lisa A Näited küsimustikust Likerti skaalal	45

Lisa B	Empiirilised näited	47
Lisa C	Programmiteegi implementatsioon	49
Lisa D	Litsents	54

Sissejuhatus

Käesolev bakalaureusetöö on ajendatud psühholoogide probleemist – kuidas hinnata nende töövaldkonnas tihti kasutatavate valikvastustega küsimustike usaldusväärsust.

Sellest lähtuvalt on töö peamiseks eesmärgiks uurida valikvastustega küsimustike hindamiseks lihtsasti kasutatavaid kvantitatiivseid meetodeid. Lisaks sellele loodab autor pakkuda eestikeelset ülevaadet ühest valikvastustega testidega seotud matemaatiliselt raamistikust. Olemasolevad käsitlused on pigem praktilised abivahendid psühholoogidele.

Töö esimeses osas esitame vajaliku taustinfo, definitsioonid ja probleemipüstituse. Seejärel ehitame üles meile vajamineva osa probleemi vaatlemiseks sobivast matemaatilisest teooriast – defineerime testi usaldusväärsuse ning rajame vundamendi selle hindamiseks. Järgnevalt tuletame mõningad erinevad usaldusväärsuse hinnangud. Seejärel tutvustame autori loodud programmeerimiseki, mis sisaldab eelnevalt tutvustatud hinnangute implementatsioone. Töö viimases osas pakub autor välja alternatiivse lähenemise küsimustiku vastuste arvulisele tõlgendamisele.

Autori eesmärgiks on, et antud tekst oleks lihtsasti järgitav keskmisele bakalaureuseõppe läbinud matemaatikatudengile – meeldetuletuseks on valdkonnaspetsiifilised matemaatilised definitsioonid ja tulemused töö käigus jooksvalt ära toodud.

Lisaks loodab autor, et loodud programmeerimiseki pakub huvi ka psühholoogidele.

Autor soovib tänada Margus Niitsood konstruktiivse juhendamise eest ning Timo Aavat ja Kadi Kähär-Petersoni abi eest käesolevas bakalaureusetöös kasutatud artiklite kättesaadavaks tegemisel.

1 Taustinfo ja probleemipüstitus

Järgnevalt anname probleemi mõtestamiseks vajalikud taustateadmised ning peatüki lõpetuseks püstitame käesoleva töö keskse probleemi.

1.1 Küsimustik, küsimus ja Likerti skaala

Käesolev uurimus tegeleb küsimustikega (*Likert scale*), milles soovitakse hinnanguid teatud arvule küsimustele (*Likert item*) n -pallisel Likerti skaalal [7], kus n jääb enamasi kahe ja kümne vahele. Käsitleme Likerti skaalasid, mis on sümmeetrilised, see tähendab, et positiivsete ja negatiivsete vastusevariantide arv on sama. Näiteks:^①

Käesoleva bakalaureusetöö ülesehitus on loogiline.				
Ei nõustu	Ei nõustu osaliselt	Nii ja naa	Nõustun osaliselt	Nõustun

Joonis 1: Näide väitest, millele palutakse hinnangut Likerti skaalal

1.2 Likerti skaala tõlgendamine intervallskaalana

Likerti skaala tõlgendamisel intervallskaalana on välja kujunenud tava seada valikvastustele vastavusse järjestatud täisarvud, kusjuures mida positiivsem on vastuse variant, seda suurem on temale vastavusse seatud arv. Reeglina kasutatakse kas arve alates ühest kuni valikvastuste arvuni või valitakse välja täisarvud nii, et neutraalsele vastusevariandile vastab null.

Käesoleva bakalaureusetöö ülesehitus on loogiline.				
-2	-1	0	1	2
↑	↑	↑	↑	↑
Ei nõustu	Ei nõustu osaliselt	Nii ja naa	Nõustun osaliselt	Nõustun
↓	↓	↓	↓	↓
1	2	3	4	5

Joonis 2: Näide kahest levinumast Likerti skaala tõlgendusest intervallskaalana

^①Terviklike küsimustike näited on esitatud lisades, joonisel 14 ja 15

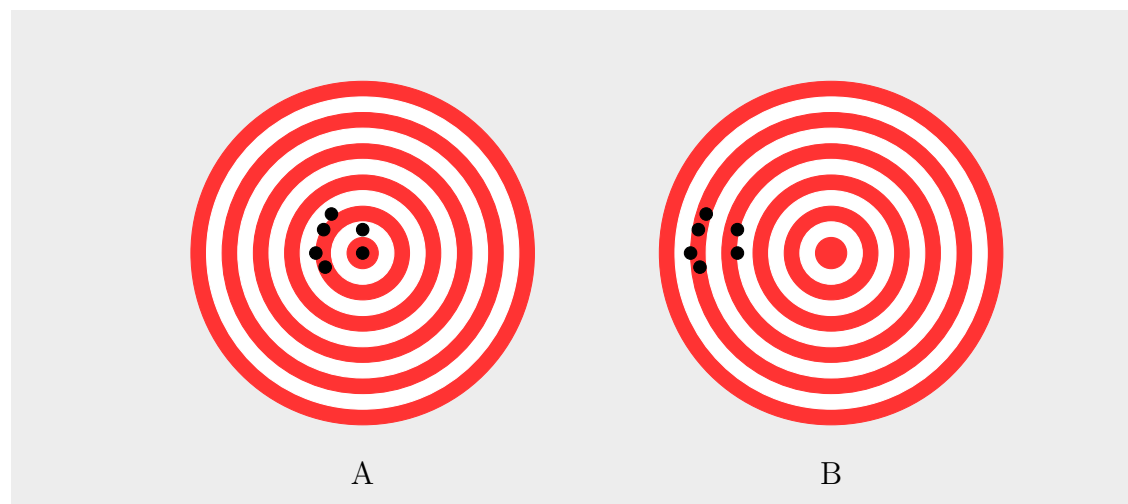
Lugejal võib tekkida õigustatud küsimus, kuidas põhjendab autor Likerti skaala käsitlemist intervallskaalana, kui Likerti skaala on olemuselt järjestik skaala ning selle tõlgendamine intervallskaalana on vastuoluline küsimus, näiteks [11]. Siinkohal tõdeme, et Likerti skaala tõlgendamine intervallskaalana on praktikas piisavalt levinud, et selle valdkonna uurimine oleks õigustatud, olenemata selle teoreetilisest põhjendatusest. Siinkohal väärrib autori silmis esile toomist kriitikute üks levinumaid argumente: „hea“ ja „väga hea“ keskmine ei ole loomulikul viisil tõlgendatav kui „hea + pool“ ehk pole põhjust eeldada, et kõikide küsimuste omavaheline kaugus on mingil põhjusel võrdne.

1.3 Küsimustiku valiidsus

Küsimustiku **valiidsus** on küsimustiku karakteristik, mis iseloomustab küsimustiku võimet mõõta seda, mida see kujundati mõõtma. Enamikes käsitlustes vaadeldakse valiidsust kui väärtust intervallskaalal nulli ja ühe vahel.

Näitena olgu meil küsimustikuks kaal, mille näitu vaatleme kui küsimustiku tulemust kaalutava isiku puhul. Kaal, mis näitab 75 kilo kaaluva inimese kaaluks 74,5 kilo, omab kõrgemat valiidsust, kui kaal, mis sama inimese puhul näitab kaaluks 65 kilo.

Küsimustiku valiidsuse mõtestamisel võime kasutada analoogiat täpsuslaskmisega. Kui vaatleme küsimustikke kui laskureid, siis tähendab see, et laskuri *A* valiidsus on kõrgem kui laskuri *B* oma seda, et tema tabamused asuvad märklaua keskkohale lähemal. Olukorda illustreerib järgnev joonis:



Joonis 3: Kahe küsimustiku valiidsuse võrdlus

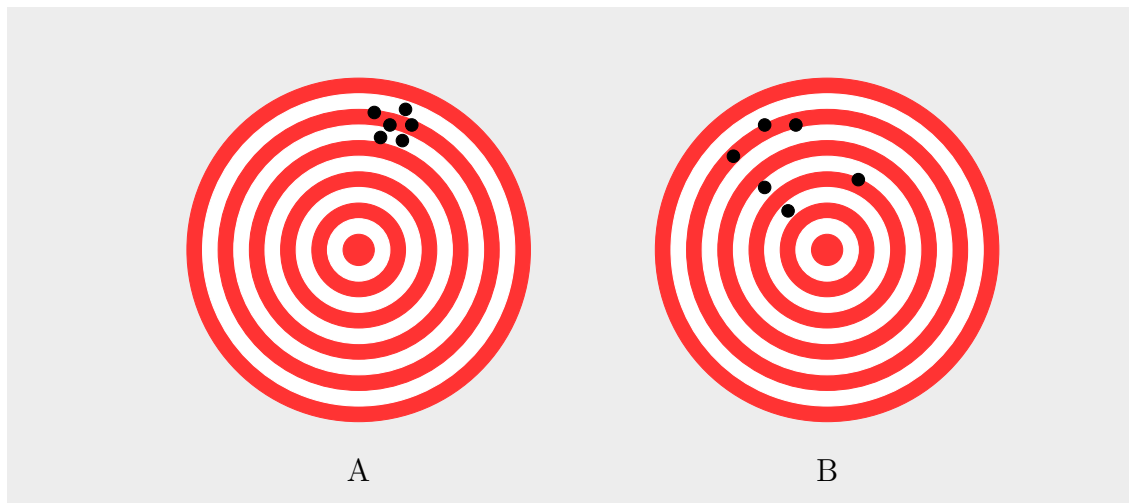
Arusaadavatel põhjustel on testi valiidsus äärmiselt oluline ning psühholoogiliste testide valiidsuse hindamine on olnud üks psühhomeetria põhilistest uurimisobjektidest. Sellega seoses on tehtud palju tööd ka valiidsuse definitsiooni täpsemaks muutmiseks, võrreldes selle paragrahvi alguses antud väga intuiitivse definitsiooniga. Kuna valiidsus on käesoleva töö raames vajalik vaid taustinfona probleemi mõtestamiseks, tundub antud lihtne definitsioon aga sobivaim.

1.4 Küsimustiku reliaablus

Küsimustiku **reliaablus** on küsimustiku karakteristik, mis kirjeldab, kui järjepidevad on subjekti erinevad mõõtmised sama küsimustikuga. Nii nagu valiidsus, on ka reliaablus reeglina määratud intervallskaalal nulli ja ühe vahel.

Näiteks kaal, mis näitab igal kaalumisel 75 kilo kaaluva inimese kaaluks 74,5 kilo, omab kõrgemat reliaablust, kui kaal, mis näitab juhuslikult kas 74,5 kilo või 75,5 kilo.

Ka küsimustiku reliaabluse mõtestamisel saame kasutada analoogiat täpsuslaskmisega. Sellisel juhul tähendab see, et laskuri *A* reliaablus on suurem kui laskuri *B* oma seda, et laskuri *A* tabamused asuvad teineteisele lähemal, kui laskuri *B* tabamused. Olukorda illustreerib järgnev joonis:



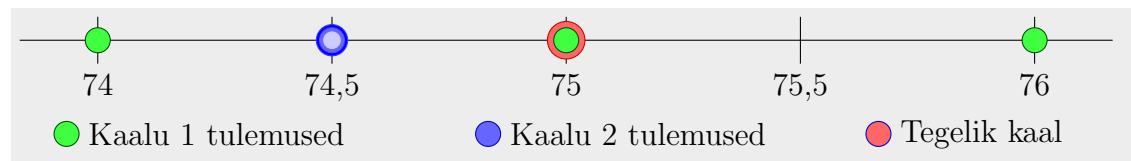
Joonis 4: Kahe küsimustiku reliaabluse võrdlus

Paneme tähele, et kõrgem reliaablus ei tähenda tingimata seda, et testi valiidsus on suurem.

Reliaabluse täpsema, matemaatilise definitsiooni anname edaspidi.

1.5 Valiidsuse ja reliaabluse suhe

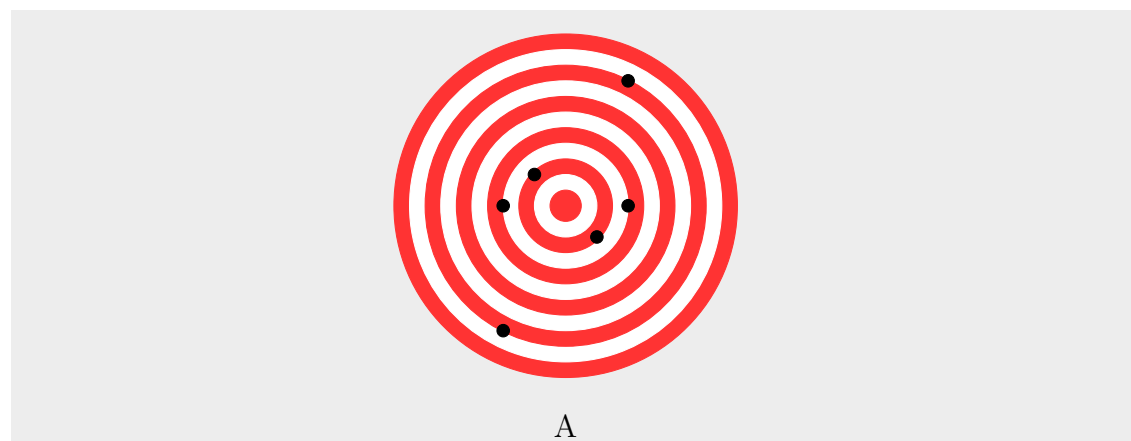
Valiidsuse ja reliaabluse suhestamisel kerkib kiiresti üles loomulik küsimus – kas test võib olla samaaegselt suure valiidsuse ja väikse reliaablusega? Eelpool antud definitsioonid jätavad selle küsimuse lahtiseks. Vaatleme järgmist olukorda:



Joonis 5: Kaalumiste tulemused kahe erineva kaaluga kolmel katsel

Kui mõista valiidsust kui küsimustiku keskmist tulemust, siis võime väita, et esimene kaal on väiksema reliaabluse, aga suurema valiidsusega. Selline valiidsuse mõtestamine ei ole aga küsimustike modelleerimise korral otstarbekas – katsete kordamine on enamasti keeruline ning väheste mõõtmistulemuste põhjal ei ole võimalik tegeliku väikese reliaablusega testi tulemust välja selgitada.

Alternatiivselt võime ka siin vaadelda olukorda täpsuslaskmise analoogiat kasutades. Olgu meil laskur A , kelle lasud paiknevad märklaua erinevates punktides, kuid nende geomeetriline keskpunkt asub märklaua keskel. Intuitiivselt on selge, et kui laskude arv on piiratud, ei ole selline definitsioon, mille korral laskuri A valiidsus oleks suur, otstarbekas. Kirjeldatud olukorda illustreerib järgnev joonis:



Joonis 6: Küsimistik, mille reliaablus on väike ja „valiidsus“ suur

Seega, meie vaadeldava olukorra – psühholoogiliste küsimustike – puhul on mingi valiidsuse taseme jaoks tarvilik tingimus mingi reliaabluse tase.

1.6 Sisemine järjepidevus

Reliaabluse kui termini probleemiks on tema mitmetähenduslikkus. Toome siinkohal ära ühe mõiste, mida tihti samuti reliaablusena või sisemise järjepidevuse reliaablusena tuntakse.

Sisemine järjepidavus (*internal consistency*) on küsimustiku karakteristik, mis iseloomustab küsimustiku erinevate küsimuste vastuste järjepidevust ehk seda, kui hästi on vastused ühist konstruktsiooni hindavatele küsimustele kooskõlas [9, 177]. Paneme tähele, et erinevalt eelnevalt toodud reliaabluse definitsioonist, on sisemine järjepidevus ühe küsimustiku läbiviimise keskne. Piltlikult väljendudes olgu meil järgnev küsimustik:

Käesoleva bakalaureusetöö ülesehitus on loogiline.				
Ei nõustu	Ei nõustu osaliselt	Nii ja naa	Nõustun osaliselt	Nõustun
Mulle meeldib käesoleva bakalaureusetöö ülesehitus.				
Ei nõustu	Ei nõustu osaliselt	Nii ja naa	Nõustun osaliselt	Nõustun
Käesolevat bakalaureusetööd on lihtne lugeda.				
Ei nõustu	Ei nõustu osaliselt	Nii ja naa	Nõustun osaliselt	Nõustun

Joonis 7: Küsimustik bakalaureusetöö ülesehituse kohta

Siin on mõõdetavaks konstruktsiooniks käesoleva bakalaureusetöö ülesehitus ning kõrgeks sisemiseks reliaabluseks on vajalik kolmele näites toodud küsimusele antud vastuste kooskõla.

Selleks, et sisemine järjepidavus oleks mõtestatud, peavad meil küsimustikus olema mingi kindla konstruktsiooni mõõtmiseks rohkem kui üks küsimus – vastasel juhul on iga küsimus täielikus kooskõlas kõikide sama konstruktsiooni mõõtvate küsimustega, ehk iseendaga. Rõhutame, et see nõue ei laiene eelnevalt defineeritud reliaablusele.

Sisemise järjepidavuse mõõtmiseks on võimalik kasutada sisemise järjepidavuse teste. Segadust suurendab veelgi see, et üks nendest testidest leiab kasutamist ka

meie poolt defineeritud reliaabluse hindamisel. Loodame, et nende kahe karakteristiku eristamine aitab lugejal käesolevas töös kergemini orienteeruda.

Märgime, et lisaks siintoodud kahele definitsioonile on ka teisi reliaabluse käsitlusi [3], mis aga selle töö kontekstis ei tohiks segadust tekitada ning mille äratoomist ei ole autor pidanud siinkohal vajalikus.

1.7 Probleemi püstitus

Käesoleva tööga üritame pakkuda lahendust järgmisele praktilisele probleemile: kas on võimalik leida paremat hinnangut küsimustiku reliaabluse kohta, kui seda on tihti kasutatav Guttmani λ_3 (rohkem tuntud kui Cronbachi α)?

Küsimusele vastamiseks on meil vaja rangemat lähenemist reliaablusele ning selleks vaatleme reliaablust ühe matemaatilise raamistiku – klassikalise testiteooria – kontekstis.

2 Reliaablus klassikalises testiteoorias

Järgnevas võtame aluseks Melvin Novicki klassikalise testiteooria [14] [13] ja selle tõlgenduse Klaas Sijtsma poolt. [18, 109]

2.1 Reliaabluse definitsioon

MEELDETULETUSEKS

Definitsioon. *Elementaarsündmuseks* nimetatakse juhusliku katse võimalikku tulemust. [§]

Definitsioon. *Tõenäosuseks* nimetatakse funktsiooni P , mis igale sündmusele $A \in 2^\Omega$, kus Ω on mingi juhusliku katse kõikvõimalike elementaarsündmuste hulk, seab vastavusse arvu $P(A)$, nii et on täidetud järgmised nõuded:

1. $P(A) \geq 0, \forall A \in 2^\Omega$,
2. $P(\Omega) = 1, P(\emptyset) = 0$,
3. Kui sündmused A_1, A_2, \dots on teineteist välistavad, siis

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Definitsioon. *Diskreetseks juhuslikuks suuruseks* nimetatakse funktsiooni $X : \Omega \rightarrow \mathbb{R}$, mis võtab kas lõpliku või loenduva arvu erinevaid väärtusi $x_1, x_2, \dots (x_n)$.

[§] Siin ja edaspidi toetume tõenäosusteooria definitsioonide ja tulemuste sõnastamisel Kalev Pärna õpikule [15].

Olgu meil mingi kogum I testile vastajaid ning küsimustik, milles on k küsimust, kus iga küsimus on l -pallisel Likerti skaalal. Nummerdame küsimused naturaalarvudega 1 kuni k . On loomulik eeldada, et nii vastajate arv, küsimuste arv kui ka Likerti skaala suurus on lõplikud. Me võime vaadelda testi läbiviimist kui juhuslikku katset ning kõiki erinevaid võimalike vastuste kombinatsioone küsimustikule kui elementaarsündmusi, tähistades saadud elementaarsündmuste hulga kui Ω ning elementaarsündmuse kui ω . Eeldame ka tõenäosuste $P_i : 2^\Omega \rightarrow \mathbb{R}$ olemasolu, mida tõlgendame kui testi tulemuste esinemissagededust vastaja i korral.

$\omega \backslash$ Küsimus	I	II	II
A	Nõustun osaliselt	Nii ja naa	Nõustun
B	Ei nõustu osaliselt	Nõustun	Nii ja naa
C	Ei nõustu	Nõustun osaliselt	Nii ja naa

Joonis 8: Näited elementaarsündmustest

Olgu X_i^j hulgal Ω määratud juhuslik suurus, mille võimalikke väärtusi vaatleme kui vastaja i poolt küsimusele j antud vastuse arvulist tõlgendust. On selge, et sellise tõlgenduse korral on juhusliku suuruse X_i^j muutumispiirkond lõplik - selle hulga võimsus ei ole suurem kui l ning järelikult on juhuslik suurus X_i^j diskreetne. Defineerime vastaja i testi tulemuse X_i kui

$$X_i = \sum_{j=1}^k X_i^j.$$

Paneme tähele, et X_i on samuti diskreetne juhuslik suurus – tõepoolest, kuna tegemist on diskreetsete juhuslike suuruste summaga, siis on ta funktsioon hulgast Ω hulka \mathbb{R} . Lisaks, kuna kõikide liidetavate muutumispiirkond on tõkestatud arvuga l ning liidetavaid on k , siis ei saa funktsiooni muutumispiirkonna võimsus olla suurem kui l^k .

MEELDETULETUSEKS

Definitsioon. Diskreetse juhusliku suuruse X *jaotuseks* nimetatakse paaride komplekti $(x_i, p_i), i = 1, 2, \dots$, kus x_i on juhusliku suuruse võimalik väärtus ning $p_i = P(\{\omega | X(\omega) = x_i, \omega \in \Omega\})$.

Definitsioon. Diskreetse juhusliku suuruse X , mille jaotuseks K on paaride komplekt $k_i = (x_i, p_i), i = 1, 2, \dots$, *keskväärtuseks* nimetatakse arvu

$$\epsilon X = \sum_{\{i | k_i \in K\}} x_i p_i$$

eeldusel, et see rida koondub absoluutselt.

Definitsioon. Juhusliku suuruse X *dispersiooniks* nimetatakse arvu $DX = \epsilon[(X - \epsilon X)^2]$.

Triviaalse juhu vältimiseks eeldame edaspidi, et juhuslike suuruste X_i dispersioonid

on nullist erinevad. Defineerime küsimustiku tegeliku tulemuse vastaja i jaoks kui testi tulemuse oodatava keskvaartuse ehk

$$t_i = \epsilon[X_i].^{②}$$

Kuna juhusliku suuruse X_i võimalike väärtuste arv on lõplik, on ka tema keskvaartus t_i lõplik.

Tõlgendame vastaja i testi tulemuse ja tegeliku tulemuse vahet mõõtmisveana, tähistame seda kui E_i . Seega $X_i = t_i + E_i$.

MEELDETULETUSEKS

Lause 2.1. Kui X on diskreetne juhuslik suurus, jaotusega K , kus K on paaride komplekt $k_i = (x_i, p_i), i = 1, 2, \dots$ ning c on mingi konstant, siis $\epsilon(X + c) = \epsilon(X) + c$

Tõestus.

$$\begin{aligned} \epsilon(X + c) &= \sum_{\{i|k_i \in K\}} (x_i + c)p_i = \sum_{\{i|k_i \in K\}} x_i p_i + \sum_{\{i|k_i \in K\}} c p_i = \\ &= \sum_{\{i|k_i \in K\}} x_i p_i + c \sum_{\{i|k_i \in K\}} p_i = \sum_{\{i|k_i \in K\}} x_i p_i + c = \epsilon(X) + c \end{aligned}$$

□

Lause 2.2. Juhuslikule suurusele konstandi liitmine ei muuda dispersiooni: $D(X + c) = DX$

Tõestus.

$$D(X + c) = \epsilon[X + c - \epsilon(X + c)] \stackrel{\text{Lause 2.1}}{=} \epsilon[X + c - \epsilon(X) + c] = \epsilon[X - \epsilon(X)] = DX$$

□

Paneme tähele, et eelneva põhjal on vastaja i mõõtmisvea keskvaartus võrdne nulliga. Tõepoolest, kuna tegeliku tulemuse definitsiooni põhjal $\epsilon[X_i] = t_i$, siis saame mõõtmisvea definitsiooni arvesse võttes

$$t_i = \epsilon[X_i] = \epsilon[t_i + E_i] \stackrel{\text{Lause 2.1}}{=} t_i + \epsilon[E_i] \implies \epsilon[E_i] = 0.$$

^②Tähistame keskvaartuse tähega ϵ , kuna tavapärase tähistus E omab antud käsitluses teistsugust tähendust.

Lisaks on lihtne märgata, et $D[X_i] = D[E_i]$, sest

$$D[X_i] = D[E_i + t_i] \stackrel{\text{Lause 2.2}}{=} D[E_i].$$

Eelnevates definitsioonides keskendusime testi tulemuste mõtestamisele fikseeritud vastaja korral. Kuna reeglina mõtestatakse teste praktikas mingi vastajate kogumi raames, siis keskendume ka meie edaspidi sellele.

Vaatleme elementaarsündmustena hulga $I \times \Omega$ elemente ning eeldame, et eksisteerib tõenäosus

$$P_p : 2^{I \times \Omega} \rightarrow \mathbb{R},$$

kus funktsiooni P_p tõlgendame kui isiku $i \in I$ ja testi tulemuse $\omega \in \Omega$ koosinemise sagedus.

Paneme tähele, et sellel elementaarsündmuste hulgal võime määrata diskreetse juhusliku suuruse T , nii et

$$T((i, \omega)) = t_i, i \in I.$$

Seda juhuslikku suurust tõlgendame kui küsimustiku tegelikku tulemust. Defineerime testi vaadeldava tulemuse kui diskreetse juhusliku suuruse X , nii et

$$X((i, \omega)) = X_i(\omega)$$

ja testi vea kui diskreetse juhusliku suuruse E , nii et

$$E((i, \omega)) = E_i(\omega).$$

Paneme tähele, et kehtib $X = E + T$. Tõepoolest, olgu meil suvaline paar $(i, \omega) \in I \times \Omega$, siis

$$X((i, \omega)) = X_i(\omega) \stackrel{\text{def}}{=} t_i + E_i(\omega) = T((i, \omega)) + E((i, \omega)).$$

Definitsioon. Juhusliku suuruste X ja Y *kovariatsiooniks* nimetatakse suurust

$$\text{cov}(X, Y) = \epsilon[(X - \epsilon X)(Y - \epsilon Y)].$$

Definitsioon. Juhusliku suuruste X ja Y *korrelatsioonikordajaks* nimetatakse suurust

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{DX}\sqrt{DY}}$$

Esitame nüüd meie poolt üles ehitatud mudeli omaduse, mis osutub väga kasulikus. Olgu meil kaks suvalist küsimustikku, mille vaadeldavad tulemused on vastavalt $X = T + E$ ja $X' = T' + E'$. Saab näidata, et siis kehtib

$$\text{cov}(T, E') = 0. \quad (2.3)$$

Teisisõnu, küsimustiku tegeliku tulemuse ja küsimustiku mõõtmisvea vahel ei ole korrelatsiooni. Paneme tähele, et kuna me ei kitsendanud X ja X' valikut, siis ka $\text{cov}(T, E) = 0$. See omadus muudab mudeli rakendamise palju lihtsamaks, kuna me ei pea eeldama, et meie koostatud küsimustikul on selline omadus. Selle omaduse, mille kehtimine ei ole intuiitiivne, tõestus on mittetriviaalne ning ei mahu käesoleva töö raamidesse. Rangest tõestusest huvitatutele soovitame tutvuda Donald Zimmermani mudeli ülesehitusega [25]. Lugejale võib olla huvipakkuv see, et kõnealune omadus järeldub väga üldistest Hilberti ruumides kehtivatest omadustest [24].

Küll aga peame tegema järgmise eelduse – olgu meil kaks erinevat küsimustikku, mille mõõtmisvead on vastavalt E ja E' , siis

$$\text{cov}(E, E') = 0. \quad (2.4)$$

Tehtud eeldus tähendab seda, et mõõtmisvead ei tohi olla süstemaatilised. Näiteks, jagades küsimustiku pooleks ning vaadeldes mõlemat osa kui eraldi küsimustikku, siis esimese mõõtmisviga teades ei tohiks see meile midagi öelda teise osa mõõtmisvea kohta. Paneme tähele, et see eeldus on siia maani tehtutest praktikas kõige raskemini tagatav ning on potentsiaalne ohukoht testiteooria kasutamisel mingi küsimustiku mõtestamiseks.

Lause 2.5. Juhuslike suuruste X ja Y puhul kehtib

$$D(X + Y) = D(X) + D(Y) + 2cov(X, Y)$$

Tõestus.

$$\begin{aligned} D(X + Y) &\stackrel{def}{=} \epsilon[X + Y - \epsilon(X + Y)]^2 \stackrel{\epsilon \text{ lin.}}{=} \epsilon[X - \epsilon X + Y + \epsilon Y]^2 = \\ &= \epsilon(X - \epsilon X)^2 + 2\epsilon(X - \epsilon X)(Y - \epsilon Y) + \epsilon(Y - \epsilon Y)^2 \stackrel{cov \text{ def}}{=} \\ &= D(X) + D(Y) + 2cov(X, Y) \end{aligned}$$

□

Järgnevalt tutvume paralleelse küsimustiku mõistega, mida võime tõlgendada kui sama küsimustiku sõltumatut kordusmõõtmist samades tingimustes. Kaks küsimustikku, mille vaadeldavad tulemused üle kõigi vastajate kogumi I on $X = T + E$ ja $X' = T' + E'$ ning mille vaadeldavad tulemused vastaja $i \in I$ jaoks on $X_i = t_i + E_i$ ja $X'_i = t'_i + E'_i$, on paralleelsed parajasti siis, kui kehtivad järgmised väited:

1. $t_i = t'_i \forall i \in I$,
2. $D[E] = D[E']$.

Eelnevast järeldub lihtsalt, et

$$T = T'. \quad (2.6)$$

Tõepoolest, olgu (i, ω) suvaline paar hulgast $I \times \Omega$. Siis kehtib

$$T((i, \omega)) \stackrel{def}{=} t_i = t'_i = T'((i, \omega)).$$

Seega nõuame, et kahe paralleelse küsimustiku korral vastajate tegelikud tulemused alati ühtiksid ning vea dispersioonid oleksid võrdsed – need nõudmised sobivad hästi kirjeldama sama küsimustikuga sooritatud kordusmõõtmist.

Paneme tähele, et kahe paralleelse küsimustiku, mille vaadeldavad tulemused on X ja X' , korral kehtib

$$D[X] = D[X'], \quad (2.7)$$

kuna

$$\begin{aligned} D[X] &= D[T + E] \stackrel{\text{Lause 2.5}}{=} D[T] + D[E] + 2cov(T, E) \stackrel{(2.3)}{=} D[T] + D[E] = \\ &= D[T'] + D[E'] = D[T'] + D[E'] + 2cov(T', E') = D[T' + E'] = D[X']. \end{aligned}$$

Lause 2.8. Olgu meil juhuslikud suurused X , Y ja Z . Siis

$$\text{cov}(X + Y, Z) = \text{cov}(X, Z) + \text{cov}(Y, Z).$$

Tõestus.

$$\begin{aligned} \text{cov}(X + Y, Z) &\stackrel{\text{def}}{=} \epsilon[(X + Y - \epsilon[X + Y])(Z - \epsilon[Z])] \stackrel{\epsilon \text{ lin.}}{=} \\ &= \epsilon[(X + Y - \epsilon[X] + \epsilon[Y])(Z - \epsilon[Z])] = \\ &= \epsilon[((X - \epsilon[X]) + (Y - \epsilon[Y]))(Z - \epsilon[Z])] = \\ &= \epsilon[(X - \epsilon[X])(Z - \epsilon[Z]) + (Y - \epsilon[Y])(Z - \epsilon[Z])] = \\ &= \epsilon[(X - \epsilon[X])(Z - \epsilon[Z])] + \epsilon[(Y - \epsilon[Y])(Z - \epsilon[Z])] + \\ &= \text{cov}(X, Z) + \text{cov}(Y, Z) \end{aligned}$$

□

Paralleelse küsimustiku abil saame klassikalise testiteooria raames defineerida reliaabluse. Olgu meil kaks paralleelset küsimustikku, mille vaadeldavad tulemused üle vastajate kogumi I on vastavalt $X = T + E$ ja $X' = T' + E'$. Küsimustiku, mille vaadeldav tulemus on X , reliaablus üle vastajate kogumi I on võrdne X ja X' korrelatsioonikordajaga, mida tähistame $\rho_{XX'}$, teisisõnu

$$\rho_{XX'} = \text{corr}(X, X').$$

Eelneva põhjal saame näidata, et kehtib

$$\begin{aligned} \rho_{XX'} &\stackrel{\text{def}}{=} \frac{\text{cov}(X, X')}{\sqrt{D[X] D[X']}} \stackrel{(2.7)}{=} \frac{\text{cov}(X, X')}{\sqrt{D[X] D[X]}} = \frac{\text{cov}(T + E, T' + E')}{D[X]} \stackrel{\text{Lause 2.8}}{=} \\ &= \frac{\text{cov}(T, T') + \text{cov}(T, E') + \text{cov}(E, T') + \text{cov}(E, E')}{D[X]} \stackrel{(2.3)}{=} \\ &= \frac{\text{cov}(T, T') + \text{cov}(E, E')}{D[X]} \stackrel{\text{eeldus (2.4)}}{=} \frac{\text{cov}(T, T')}{D[X]} \stackrel{(2.6)}{=} \frac{\text{cov}(T, T)}{D[X]} \stackrel{\text{def}}{=} \\ &= \frac{\epsilon[(X - \epsilon[X])(X - \epsilon[X])]}{D[X]} \stackrel{\text{def}}{=} \frac{D[T]}{D[X]}. \end{aligned} \tag{2.9}$$

Tõestatud tulemus on huvitav – ilmneb, et saame öelda midagi küsimustiku ja sellega paralleelse küsimustiku korrelatsiooni kohta, teadmata midagi paralleelse küsimustiku kohta.

Samuti märgime, et kuna $T = X - E$, siis kehtib

$$\begin{aligned} \rho_{XX'} &= \frac{D[T]}{D[X]} \stackrel{\text{Lause 2.5}}{=} \frac{D[T + E] - 2\text{cov}(T, E) - D[E]}{D[X]} \stackrel{(2.3)}{=} \\ &\stackrel{(2.3)}{=} \frac{D[T + E] - D[E]}{D[X]} = \frac{D[X] - D[E]}{D[X]} = 1 - \frac{D[E]}{D[X]}. \end{aligned} \quad (2.10)$$

Paneme tähele, kuna $D[X] = D[T] + D[E]$, siis eelneva põhjal $\rho_{XX'} \in [0, 1]$.

Eelneva põhjal on selge, et väide „küsimustiku reliaablus on suurem“ on samaväärne järgmise kolme väitega:

1. küsimustiku korrelatsioon paralleelse küsimustikuga on suurem,
2. küsimustiku tegeliku tulemuse dispersioon on suurem võrreldes küsimustiku vaadeldava tulemuse dispersiooniga,
3. küsimustiku mõõtmisvea dispersioon on väiksem võrreldes küsimustiku vaadeldava tulemuse dispersiooniga.

Märgime, et klassikalise testiteooria raames defineeritud reliaablus sobib kokku eelmises peatükis antud mitteformaalse definitsiooniga, kui võtta arvesse paralleelse küsimustiku semantilist tähendust.

Paneme tähele, et tegemist on suhteliselt nõrku eeldusi vajava teooriaga, kõige raskem on praktikas tagada süstemaatiliste vigade puudumine.

Kuna küsimustiku korduvmõõtmiste läbiviimine ei ole tihti võimalik, on kasutusele võetud erinevad meetodid testi reliaabluse hindamiseks küsimustiku ühekordse läbiviimise käigus saadud informatsiooni põhjal, kus kasutatakse punktides (2.9) ja (2.10) toodud omadusi. Järgnevalt vaatleme, kuidas need meetodid on tuletatud.

2.2 Vajalikud taustteadmised reliaabluse hindamiseks

Järgnev käsitlus tugineb Jacksoni ja Anguwamba 1977. aasta artiklile. [10] Olgu meil vastajate populatsioon I , üle mille vaatleme küsimustikku, milles on k küsimust. Vaatleme seda küsimustikku klassikalise testiteooria raames ning olgu X küsimustiku vaadeldav tulemus, T küsimustiku tegelik tulemus ning E küsimustiku mõõtmisvea. Vaatleme iga küsimust kui ühest küsimusest koosnevat alamküsimustikku ja tähistame fikseeritud küsimuse $i \in \{1, 2, \dots, k\}$ vaadeldava tulemuse kui X_i , tegeliku tulemuse kui T_i ning mõõtmisvea E_i . Seega $X = \sum_{i=0}^k X_i$,

$T = \sum_{i=0}^k T_i$ ja $E = \sum_{i=0}^k E_i$. Paneme tähele, et kui eelnevas peatükis tähistasime alaindeksiga küsimusele vastajat, siis edaspidi tähistame sellega kindlat küsimust.

Vaatleme testi küsimuste tulemuste ja mõõtmisvea vektoreid: $(X_1, X_2, X_3, \dots, X_k)$, $(T_1, T_2, T_3, \dots, T_k)$, $(E_1, E_2, E_3, \dots, E_k)$. Tähistame nende vektorite kovariatsioonimatriksid^③ vastavalt $\Sigma_X, \Sigma_T, \Sigma_E$. Seega,

$$\Sigma_X = \begin{pmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \text{cov}(X_1, X_3) & \cdots & \text{cov}(X_1, X_k) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \text{cov}(X_2, X_3) & \cdots & \text{cov}(X_2, X_k) \\ \text{cov}(X_3, X_1) & \text{cov}(X_3, X_2) & \text{cov}(X_3, X_3) & \cdots & \text{cov}(X_3, X_k) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_k, X_1) & \text{cov}(X_k, X_2) & \text{cov}(X_k, X_3) & \cdots & \text{cov}(X_k, X_k) \end{pmatrix}.$$

Eelnevalt läbiviidud arutelu (vt. (2.9)) põhjal teame, et $\text{cov}(X_i, X_j) = \text{cov}(T_i, T_j) + \text{cov}(E_j, E_j)$. Seega,

$$\Sigma_X = \Sigma_T + \Sigma_E. \quad (2.11)$$

Lisaks kehtib eelduse (2.4) põhjal, et juhul kui $i \neq j$, siis $\text{cov}(E_i, E_j) = 0$. Sellest tulenevalt kehtib

$$\Sigma_E = \begin{pmatrix} \text{cov}(E_1, E_1) & 0 & 0 & \cdots & 0 \\ 0 & \text{cov}(E_2, E_2) & 0 & \cdots & 0 \\ 0 & 0 & \text{cov}(E_3, E_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \text{cov}(E_k, E_k) \end{pmatrix}. \quad (2.12)$$

Tähistame edaspidi matriksi Σ_X elemente kui x_{ij} , matriksi Σ_T elemente kui t_{ij} ning matriksi Σ_E peadiagonaali elemente kui θ_i . Seega, kui $i \neq j$, siis $t_{ij} = x_{ij}$ ning $t_{ii} = x_{ii} - \theta_i$.

Tuletame meelde, et eelmises peatükis veendusime (2.10), et testi reliaablus võrdub avaldisega $1 - \frac{D[E]}{D[X]}$.

Veendume, et juhuslike suuruste summa dispersioon võrdub liidetavate kovariatsioonimatriksi elementide summaga, ehk kui $X = \sum_{i=1}^k X_i$ ja $\Sigma_X = (x_{ij})$ on vektori

^③ Praktikas ei ole kovariatsioonimatriksid meile teada, kuid valimi põhjal on meil võimalik saada hinnang vaadeldavate tulemuste kovariatsioonimatriksi kohta.

(X_1, X_2, \dots, X_K) kovariatsioonimaatriks, siis

$$D[X] = \sum_{i=1}^k \sum_{j=1}^k x_{ij}. \quad (2.13)$$

Eelnevas veendumiseks paneme tähele, et

$$\begin{aligned} D[X] &= D[X_1 + X_2 + \dots + X_k] \stackrel{\text{Lause 2.5}}{=} \\ &= D[X_1] + D[X_2 + X_3 + \dots + X_k] + 2\text{cov}(X_1, X_2 + X_3 + \dots + X_k) \stackrel{\text{Lause 2.8}}{=} \\ &= D[X_1] + D[X_2 + X_3 + \dots + X_k] + \\ &\quad + 2[\text{cov}(X_1, X_2) + \text{cov}(X_1, X_3) + \dots + \text{cov}(X_1, X_k)]. \end{aligned}$$

Näeme, et tegemist on rekurrentse võrrandiga. Arvestades, et $\text{cov}(X, X) = D[X]$ (vt.(2.9)), näeme, et igale liidetavale eelnevas summas vastab mingite kovariatsioonimaatriksite elementide summa, kusjuures kõik kovariatsioonimaatriksi elemendid kuuluvad täpselt ühte nendest summadest. Kovariatsiooni maatriksi jagunemist illustreerib järgnev tabel, kus üleval vasakus nurgas on $D[X_1]$, tumedama tooniga on tähistatud liidetavad, mille summa on võrdne arvuga $2[\text{cov}(X_1, X_2) + \text{cov}(X_1, X_3) + \dots + \text{cov}(X_1, X_k)]$ ja heledama tooniga liikmed, mille summast moodustub $D[X_1 + X_2 + \dots + X_k]$. Siinkohal meenutame, et kovariatsioon on sümmeetriline ning seega $\text{cov}(X_i, X_j) = \text{cov}(X_j, X_i)$.

$\text{cov}(X_1, X_1)$	$\text{cov}(X_1, X_2)$	$\text{cov}(X_1, X_3)$	\dots	$\text{cov}(X_1, X_k)$
$\text{cov}(X_2, X_1)$	$\text{cov}(X_2, X_2)$	$\text{cov}(X_2, X_3)$	\dots	$\text{cov}(X_2, X_k)$
$\text{cov}(X_3, X_1)$	$\text{cov}(X_3, X_2)$	$\text{cov}(X_3, X_3)$	\dots	$\text{cov}(X_3, X_k)$
\vdots	\vdots	\vdots	\ddots	\vdots
$\text{cov}(X_k, X_1)$	$\text{cov}(X_k, X_2)$	$\text{cov}(X_k, X_3)$	\dots	$\text{cov}(X_k, X_k)$

Lisaks on sellel rekurrentsel võrrandil lihtne baasjuht: selleks on olukord, kus summa koosneb ainult ühest liidetavast.

Arvestades maatriksi 2.12 struktuuri, saame võrduse

$$1 - \frac{D[E]}{D[X]} = 1 - \frac{\sum_i \theta_i}{\sum_i \sum_j x_{ij}}. \quad (2.14)$$

MEELDETULETUSEKS

Definitsioon. Maatriks $C \in \text{Mat}_k(\mathbb{R})$ on *mittenegatiivselt määratud*, kui iga maatriksi $v \in \text{Mat}_{k,1}(\mathbb{R})$ korral $v^T C v \geq 0$.

Veendume järgnevalt, et iga kovariatsioonimaatriks ning järelikult ka Σ_X, Σ_T ja Σ_E on mittenegatiivselt määratud.

MEELDETULETUSEKS

Lause 2.15. Kui X ja Y on juhuslikud suurused ja $\alpha \in \mathbb{R}$, siis

$$\text{cov}(X, \alpha Y) = \alpha \text{cov}(X, Y)$$

Tõestus.

$$\begin{aligned} \text{cov}(X, \alpha Y) &= \epsilon[(X - \epsilon[X])(\alpha Y - \epsilon[\alpha Y])] \stackrel{\epsilon \text{ lin.}}{=} \\ &= \epsilon[(X - \epsilon[X])(\alpha Y - \alpha \epsilon[Y])] = \epsilon[\alpha(X - \epsilon[X])(Y - \epsilon[Y])] = \\ &= \alpha \epsilon[(X - \epsilon[X])(Y - \epsilon[Y])] = \alpha \text{cov}(X, Y) \end{aligned}$$

□

Olgu meil suvaliste juhuslike suuruste vektor (Y_1, Y_2, \dots, Y_n) , kus $n \in \mathbb{N}$ ja olgu Σ_Y selle vektori kovariatsioonimaatriks. Olgu meil suvaline n -vektor $v = (v_1, v_2, \dots, v_n)$. Konstrueerime uue juhuslike suuruste vektori $(Y'_1, Y'_2, \dots, Y'_n)$, kusjuures $Y'_i = v_i Y_i$ ning olgu Σ'_Y selle vektori kovariatsioonimaatriks. Vaatleme juhuslikku suurust $Y' = \sum_{i=1}^n Y'_i$. Kuna eelneva põhjal $D[Y'] = \sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Y_{ij}}$ ja definitsiooni põhjal $D[Y'] \geq 0$, siis $\sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Y_{ij}} \geq 0$. Kuna Σ'_Y on kovariatsioonimaatriks, siis

$$\sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Y_{ij}} = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(v_i Y_i, v_j Y_j) \stackrel{\text{Lause 2.15}}{=} \sum_{i=1}^n \sum_{j=1}^n v_i v_j \text{cov}(Y_i, Y_j). \quad (2.16)$$

Paneme tähele, et

$$v^T \Sigma_Y v =$$

$$\begin{aligned}
&= \begin{pmatrix} v_1 \text{cov}(Y_1, Y_1) + v_2 \text{cov}(Y_1, Y_2) + \dots v_n \text{cov}(Y_1, Y_n) \\ v_1 \text{cov}(Y_2, Y_1) + v_2 \text{cov}(Y_2, Y_2) + \dots v_n \text{cov}(Y_2, Y_n) \\ \vdots \\ v_1 \text{cov}(Y_n, Y_1) + v_2 \text{cov}(Y_n, Y_2) + \dots v_n \text{cov}(Y_n, Y_n) \end{pmatrix} v = \\
&\quad \sum_{i=1}^n \sum_{j=1}^n v_i v_j \text{cov}(Y_i, Y_j) \stackrel{(2.16)}{=} \sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Yij} = D[Y']
\end{aligned}$$

Seega $v^T \Sigma_Y v \geq 0$, ning kuna v oli valitud suvaliselt, siis Σ_Y on mittenegatiivselt määratud. Kuna n ja Y_1, Y_2, \dots, Y_n olid valitud suvaliselt, siis on iga kovariatsioonimaatriks mittenegatiivselt määratud.

Sellega oleme andnud reliaabluse matemaatilise definitsiooni ning mõtestanud lahti vajaminevad taustteadmised reliaabluse hindamiseks. Järgnevalt tuletame mõned reliaabluse alumised tõkked.

3 Reliaabluse alumised tõkked

Järgnevalt vaatleme reliaabluse viite alumist tõket ning üritame pakkuda lahenduse töö alguses püstitatud probleemile. Vaadeldavatest alumistest tõketest neli – $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ – pakkus 1945. aastal välja Louis Guttman [8]. Viienda vaadeldava tõkke pakkusid välja Jackson ja Anguwamba oma 1977. aasta artiklis [10]. Kõik tõkked tuletame Jacksoni ja Anguwamba artiklile tuginedes.

Olgu meil vastajate kogum I , üle mille vaatleme k küsimusest koosnevat küsimustikku, mille vaadeldav tulemus on X , tegelik tulemus on T ja mõõtmisviga on E . Nummerdame küsimused arvudega 1 kuni k . Vaatleme iga küsimust eraldi alamküsimustikuna, ning olgu X_i küsimuse i vaadeldav tulemus, T_i küsimuse i tegelik tulemus ja E_i küsimuse i mõõtmisviga. Analoogiliselt eelmise peatükiga tähistame vektorite (X_1, X_2, \dots, X_k) , (T_1, T_2, \dots, T_k) ja (E_1, E_2, \dots, E_k) kovariatsioonimaatrikseid kui $\Sigma_X, \Sigma_T, \Sigma_E \in \text{Mat}_k(\mathbb{R})$. Nagu eelmises peatükis, tähistame ka siin maatriksi Σ_X elemente kui x_{ij} , maatriksi Σ_T elemente kui t_{ij} ning maatriksi Σ_E peadiagonaali elemente kui θ_i . Edaspidi vaatleme selles peatükis küsimustiku, mille vaadeldav tulemus on X , reliaabluse alumisi tõkkeid.

3.1 Alumine tõke λ_1

Vaatleme kovariatsioonimaatriksit $\Sigma_T = (t_{ij})$. Eeldame vastuväiteliselt, et maatriksi Σ_T peadiagonaalil leidub element, mis on negatiivne. Olgu selleks elemendiks t_{ii} . Olgu v vektor, millel on k elementi, kusjuures

$$v_j = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}.$$

Paneme tähele, et nüüd $v^T \Sigma_T v = t_{ii} < 0$. Tõepoolest,

$$v^T \Sigma_T v = (t_{i1} \ t_{i2} \ \dots \ t_{ik}) v = t_{ii}.$$

Seega ei ole maatriks enam positiivselt määratud, mis on eelmises peatükis läbi viidud arutelu põhjal vastuolus sellega, et Σ_T on kovariatsioonimaatriks.

Seega iga i korral peab kehtima $t_{ii} \geq 0$ ning kuna eelmise peatüki põhjal $\Sigma_X = \Sigma_T + \Sigma_E$, siis peab kehtima $\theta_i < x_{ii}$ iga $i, 1 \leq i \leq k$ korral.

Seega ka $\sum_{i=1}^k \theta_i \leq \sum_{i=1}^k x_{ii}$. Järelikult ka

$$1 - \frac{\sum_{i=1}^k x_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \leq 1 - \frac{\sum_{i=1}^k \theta_i}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{D[E]}{D[X]}. \quad (3.1)$$

Seega, võttes arvesse võrdust (2.14), oleme tõestanud järgmise lemma:

Lemma 1. Küsimustiku, mille vaadeldav tulemus on X , reliaabluse alumiseks tõkkeks on λ_1 , kus

$$\lambda_1 = 1 - \frac{\sum_{i=1}^k x_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}}.$$

3.2 Alumine tõke λ_2

Enne järgmise alumise tõkke tuletamist veendume, et kehtivad kaks abitulemust.

Lemma 2. Olgu meil mitte-negatiivselt määratud sümmeetriline maatriks $C = (c_{ij})$, $C \in \text{Mat}_n(\mathbb{R})$. Siis kehtib iga i, j korral

$$c_{ii}c_{jj} \geq c_{ij}^2.$$

Tõestus. Valime suvaliselt i ja j , nii et $i \neq j$, $1 \leq i, j \leq n$ ning $\alpha \in \mathbb{R}$. Olgu meil n -vektor v , kus

$$v_k = \begin{cases} 1, & k = i \\ \alpha, & k = j \\ 0, & k \notin \{i, j\} \end{cases}.$$

Kuna eelduste kohaselt on C mitte-negatiivselt määratud, siis peab kehtima

$$0 \leq v^T C v = \begin{pmatrix} \mathbf{0}^T & 1 & \mathbf{0}'^T & \alpha & \mathbf{0}''^T \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \\ \mathbf{0}' \\ \alpha \\ \mathbf{0}'' \end{pmatrix}, \quad (3.2)$$

kus $\mathbf{0}, \mathbf{0}', \mathbf{0}''$ on nullvektorid pikkustega vastavalt $i-1, j-i-1, n-j$.

Arvutame avaldise 3.2 väärtuse:

$$\begin{aligned}
v^T C v &= \begin{pmatrix} \mathbf{0}^T & 1 & \mathbf{0}'^T & \alpha & \mathbf{0}''^T \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \\ \mathbf{0}' \\ \alpha \\ \mathbf{0}'' \end{pmatrix} = \\
&= \begin{pmatrix} \mathbf{0}^T & 1 & \mathbf{0}'^T & \alpha & \mathbf{0}''^T \end{pmatrix} \begin{pmatrix} c_{1i} + \alpha c_{1j} \\ \vdots \\ a_{ni} + \alpha c_{nj} \end{pmatrix} = \\
&= c_{ii} + \alpha c_{ij} + \alpha c_{ji} + \alpha^2 c_{jj} = c_{ii} + 2\alpha c_{ij} + \alpha^2 c_{jj}.
\end{aligned}$$

Seega saame, et $0 \leq c_{ii} + 2\alpha c_{ij} + c_{jj}$. See, et saadud ruutvõrratus kehtib, on samaväärne sellega, et vaadeldava ruutvõrrandi diskriminant ei ole nullist suurem. Seega,

$$c_{ij}^2 - c_i c_j \leq 0,$$

milles soovisimegi veenduda. ^④. □

Lemma 3. Olgu meil mingid arvud $y_i, 1 \leq i \leq n$. Siis kehtib järgnev võrdus

$$2(n-1) \left(\sum_{i=1}^n y_i \right)^2 = \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (y_i - y_j)^2 + 2n \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n y_i y_j.$$

Tõestus. Paneme tähele, et kehtib järgnev võrdus

$$\left(\sum_{i=1}^n y_i \right)^2 = \sum_{i=1}^n \sum_{j=1}^n y_i y_j = \sum_{i=1}^n y_i^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n y_i y_j.$$

Järelikult kehtib ka võrdus

$$2(n-1) \left(\sum_{i=1}^n y_i \right)^2 = 2(n-1) \sum_{i=1}^n y_i^2 + 2(n-1) \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n y_i y_j.$$

^④ Eelneva arutelu idee pärineb Ivar Tammeraidi õpikust. [21]

Liidetavad teistmoodi grupeerides saame, et

$$\begin{aligned}
& 2(n-1) \sum_{i=1}^n y_i^2 + 2(n-1) \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n y_i y_j = \\
& = (y_1^2 - 2y_1 y_2 + y_2^2) + (y_1^2 - 2y_1 y_3 + y_3^2) + \cdots + (y_1^2 - 2y_1 y_n + y_n^2) + \\
& + (y_2^2 - 2y_2 y_1 + y_1^2) + (y_2^2 - 2y_2 y_3 + y_3^2) + \cdots + (y_2^2 - 2y_2 y_n + y_n^2) + \\
& + \cdots + \\
& + (y_n^2 - 2y_n y_1 + y_1^2) + (y_n^2 - 2y_n y_2 + y_2^2) + \cdots + (y_n^2 - 2y_n y_{n-1} + y_{n-1}^2) + \\
& + 2n \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n y_i y_j = \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (y_i - y_j)^2 + 2n \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n y_i y_j.
\end{aligned}$$

Selles aga soovisimegi veenduda.

□

Vaatleme jällegi kovariatsiooni maatriksit $\Sigma_E = \text{diag}(\theta_i)$. Märkame, et eelmise peatüki põhjal kehtib $x_{ii} = \theta_i + t_{ii}$ iga $1 \leq i \leq k$ korral. Siis kehtib ka järgnev:

$$\sum_{i=1}^k \theta_i = \sum_{i=1}^k x_{ii} - \sum_{i=1}^k t_{ii}.$$

Märgime, et kuna kovariatsioon on definitsiooni põhjal sümmeetriline, siis on ka suvaliste juhuslike suuruste vektori (Y_1, Y_2, \dots, Y_n) kovariatsioonimaatriks Σ_Y sümmeetriline. Tõepoolest, iga $1 \leq i, j \leq n$ korral

$$\Sigma_{Y_{ij}} = \text{cov}(Y_i, Y_j) = \text{cov}(Y_j, Y_i) = \Sigma_{Y_{ji}}.$$

Arvestades kovariatsioonimaatriksite mitte-negatiivselt määratust ja sümmeetrilisust saame, et

$$\sum_{i=1}^k t_{ii} \stackrel{\text{Lemma 3}}{=} \sqrt{\frac{\sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k (t_{ii} - t_{jj})^2 + 2k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^{i-1} t_{ii} t_{jj}}{2(k-1)}} \geq \sqrt{\frac{k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k t_{ii} t_{jj}}{(k-1)}}.$$

Eelnevalt tõestatud abitulemuse ja selle, et kui $i \neq j$, siis $t_{ij} = x_{ij}$, põhjal saame väita, et

$$\sum_{i=1}^k t_{ii} \geq \sqrt{\frac{k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k t_{ii} t_{jj}}{(k-1)}} \stackrel{\text{Lemma 2}}{\geq} \sqrt{\frac{k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k t_{ij}^2}{(k-1)}} = \sqrt{\frac{k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}^2}{(k-1)}}$$

Paneme tähele, et eelneva põhjal kehtib:

$$1 - \frac{\sum_{i=1}^k x_{ii} - \sqrt{\frac{k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}^2}{(k-1)}}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \leq 1 - \frac{\sum_{i=1}^k x_{ii} - \sum_{i=1}^k t_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{\sum_{i=1}^k \theta_i}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{D[E]}{D[X]}.$$

Seega kehtib järgnev lemma.

Lemma 4. Küsimustiku, mille vaadeldav tulemus on X , reliaabluse alumiseks tõkkeks on λ_2 , kus

$$\lambda_2 = 1 - \frac{\sum_{i=1}^k x_{ii} - \sqrt{\frac{k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}^2}{(k-1)}}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}}.$$

Paneme tähele, et

$$\lambda_2 = \lambda_1 + \frac{\sqrt{\frac{k \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}^2}{(k-1)}}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}}.$$

ning seega $\lambda_1 \leq \lambda_2$.

3.3 Alumine tõke λ_3

Vaatleme jällegi kovariatsioonimaatriksit Σ_T . Olgu meil arvud i, j , nii et $1 \leq i, j \leq k, j \neq i$ ning olgu meil k -vektor v , mille element kohal i on 1 ja element kohal j on -1 . Kuna kovariatsioonimaatriks peab olema mitte-negatiivselt määratud, siis peab kehtima

$$0 \leq v^T \Sigma_T v = v^T \begin{pmatrix} t_{1i} - t_{1j} \\ t_{2i} - t_{2j} \\ \vdots \\ t_{ki} - t_{kj} \end{pmatrix} = t_{ii} + t_{jj} - 2t_{ij}.$$

Võttes arvesse, et kui $i \neq j$, siis $x_{ij} = t_{ij}$, saame, et $t_{ii} + t_{jj} \geq 2x_{ij}$.

Teame, et võimalusi valida i ja j nii, et $i \neq j$ on $k(k-1)$ tükki. Summeerides kõikide võimalike i ja j valikul saadud võrratused saame, et

$$2(k-1) \sum_{i=1}^k t_{ii} \geq 2 \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij},$$

mis on samaväärne sellega, et

$$\sum_{i=1}^k t_{ii} \geq \frac{1}{k-1} \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}.$$

Märgime, et järelikult oleme leidnud järjekordse reliaabluse alumise tõkke. Tõepoolest,

$$1 - \frac{\sum_{i=1}^k x_{ii} - \frac{1}{k-1} \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \leq 1 - \frac{\sum_{i=1}^k x_{ii} - \sum_{i=1}^k t_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{\sum_{i=1}^k \theta_i}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{D[E]}{D[X]}.$$

Sõnastame saadud tulemuse lemmmana.

Lemma 5. Küsimustiku, mille vaadeldav tulemus on X , reliaabluse alumiseks tükkeks on λ_3 , kus

$$\lambda_3 = 1 - \frac{\sum_{i=1}^k x_{ii} - \frac{1}{k-1} \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}}.$$

Tükke teisendamisel saame, et

$$\begin{aligned} \lambda_3 &= 1 - \frac{\sum_{i=1}^k x_{ii} - \frac{1}{k-1} \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = \frac{\sum_{i=1}^k \sum_{j=1}^k x_{ij} - \sum_{i=1}^k x_{ii} + \frac{1}{k-1} \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = \\ &= \frac{\sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij} + \frac{1}{k-1} \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = \frac{k}{k-1} \left(\frac{\sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k x_{ij}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \right) = \\ &= \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k x_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \right). \end{aligned}$$

Tuletades meelde tükke λ_1 definitsiooni (lemma ??) on ilmne, et

$$\lambda_3 = \frac{k \lambda_1}{k-1} \quad (3.3)$$

ja seega $\lambda_1 \leq \lambda_3$. Lisaks on võimalik näidata, et $\lambda_3 \leq \lambda_2$, näiteks [17, 148-149].

Reliaabluse alumine tõke λ_3 on enam tuntud kui Cronbachi α , mis on nimetatud Lee Cronbachi järgi, kes kasutas seda oma 1951. aasta artiklis [4], ning see on ilmselt enim kasutatud reliaabluse hinnang. Käesolevas töös jääme selle alumise tükke nimetamisel siiski λ_3 juurde – nii nimetas tuletatud tüket Louis Guttman oma kuus aastat varem ilmunud artiklis [8].

Siinkohal märgime ka ära, et λ_3 kasutatakse tihti ka testi sisemise järjepidavuse hinnanguks ning sellega kokku puutudes tuleb kindlasti välja selgitada, mida autor reliaabluse all mõistab.

3.4 Alumine tõke λ_4

Vaatleme kovariatsioonimaatriksit $\Sigma_T = (t_{ij})$. Vaatleme k vektorite hulka $V = \{v : |v_i| = 1 \ \forall i, 1 \leq i \leq k\}$. Olgu meil suvaline vektor $v \in V$. Kuna Σ_T on mitte-negatiivselt määratud, siis teame, et kehtib

$$0 \leq v^T \Sigma_T v.$$

Arvestades kovariatsioonimaatriksi $\Sigma_E = \text{diag}(\theta_i)$ omadusi, saame võrduse (2.11) põhjal

$$\sum_{i=1}^k \theta_i = \sum_{i=1}^k v_i^2 \theta_i = v^T \Sigma_E v \leq v^T \Sigma_X v.$$

Seega, iga $v \in V$ korral saame reliaabluse alumise tõkke. Tõepoolest, (2.14) põhjal iga vektori $v \in V$ korral

$$1 - \frac{v^T \Sigma_X v}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \leq 1 - \frac{\sum_{i=1}^k \theta_i}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{D[E]}{D[X]}.$$

Defineerime uue alumise tõkke kui maksimumi kõikidest sellistest tõketest.

Lemma 6. Küsimustiku, mille vaadeldav tulemus on X , reliaabluse alumiseks tõkkeks on λ_4 , kus

$$\lambda_4 = \max_{v \in V} \left(1 - \frac{v^T \Sigma_X v}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \right).$$

3.5 Alumine tõke glb

Paneme tähele, et $\Sigma_E = \text{diag}(\theta_i)$ struktuuri (2.12) tõttu on tema mitte-negatiivselt määratus samaväärne sellega, et $0 \leq \theta_i$ iga $1 \leq i \leq k$ korral – selle tulemuse tõestus on analoogiline alumise tõkke λ_1 tuletamisel läbitehtuga.

Üritame nüüd leida vea suurimat võimalikku dispersiooni, mis on (2.10) põhjal samaväärne summaga $\sum_{i=1}^k \theta_i$. Arvestades eelnevalt kirjeldatud piiranguid, saame selle kirja panna järgneva optimeerimisülesandena:

$$\begin{aligned} \max \quad & \sum_i \theta_i \\ & \theta_i > 0 \\ \Sigma_X = \Sigma_T + \Sigma_E \\ & \Sigma_T \succeq 0, \end{aligned} \tag{3.4}$$

kus kirja pilt $\Sigma_T \succeq 0$ tähistab tingimust " Σ_T on mitte-negatiivselt määratud".

Olgu z eelneva optimeerimisülesande lahend Σ_X korral. Oleme saanud küsimustiku reliaabluse järgmise alumise tõkke.

Lemma 7. Küsimustiku, mille vaadeldav tulemus on X , reliaabluse alumiseks tõkkeks on *glb* (*greatest lower bound*), kus

$$glb = 1 - \frac{z}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}}.$$

Paneme tähele, et kuna z on maksimaalne vea dispersioon fikseeritud Σ_X korral, siis on alumine tõke *glb* minimaalne võimalik reliaablus. Siit aga saame, et iga teine reliaabluse alumine tõke peab olema tõkkest *glb* väiksem või sellega võrdne. Seega on *glb* suurim võimalik alumine tõke.

3.6 Hinnangute kasutatavus praktikas

Püstitasime töö alguses küsimuse, kas on võimalik leida paremat reliaabluse hinnangut kui λ_3 . Käsitletud hinnangutest on nii *glb* kui λ_2 reliaabluse alumised tõkked ning samas vähemalt sama suured kui λ_3 , seega on need ka paremad reliaabluse hinnangud kui λ_3 . Empiirilisel on ka λ_4 suurel enamikel juhtudest suurem kui λ_3 .

Paneme tähele, et λ_3 leidmine on väga lihtne, ning eelmise sajandi keskel, mil arvutuslik võimsus oli tagashoidlik, oli tegemist olulise eelisega. Tänapäeval on λ_3 lihtsus peaaegu olematu eelis, arvestades, et ka väga suurte küsimustike puhul ei valmista paremate hinnangute leidmine optimiseeritud programmi korral probleeme.

Lisas B esitatud näidete põhjal näeme, et praktikas võivad hinnangud olla väga erinevad – näiteks joonisel 16 kujutatud kovariatsioonimaatriksi korral on tulemused järgnevad:

Hinnang	λ_1	λ_2	λ_3	λ_4	glb
Väärtus	0.39	0.60	0.43	0.73	0.77

Joonis 9: Joonisel 16 kujutatud kovariatsioonimaatriksi reliaabluse hinnangud

On lihtne märgata, et võrreldes tõketega λ_4 ja glb alahindab λ_3 reliaablust oluliselt. Samuti on märgatav vahe hinnangutel λ_2 ja λ_3 . Tuletades meelde hinnangute definitsioonid, võime üritada põhjendada hinnangute erinevust – hinnangud kasutavad kovariatsioonimaatriksitest erineva hulga informatsiooni. Kui λ_3 kasutab ainult peadiagonaalil asetsevat informatsiooni, siis tõke glb kasutab kogu kovariatsioonimaatriksis leiduvat teavet.

Joonisel 17 on aga näha, et isegi juhul, kui λ_3 on suur, annavad nii glb , λ_4 kui ka λ_3 suuremaid hinnanguid. Seega, kui küsimustikku mõtestada klassikalise testiteooria raamistikus, siis ei näe autor põhjust, miks peaks kasutama reliaabluse hindamiseks λ_3 , ning soovitab eelistada teisi, kõrgemaid hinnanguid. Autor märgib siinkohal, et tegemist tundub olevat valdkonna ekspertide seas levinud arvamusega. [17] [19] [5]

Mainime siinkohal ära, et hinnangud glb ja λ_4 võivad teatud tingimustes olla nihkega. See on probleem, mis esineb praktikas, kui kasutame kovariatsioonimaatriksi asemel tema valimi põhjal saadud hinnangut. Tõkke glb korral vaata [22], λ_4 korral [2].

4 Programmiteek

Järgnevalt põhjendan lühidalt Pythoni eeliseid teegi realiseerimisel võrreldes seliste alternatiividega nagu Java, R või MATLAB:

1. Python on keel, millega programmeerimisalased õpingud tavaliselt algavad. Näiteks Tartu Ülikoolis õpetatakse programmeerimist keeles Python nii mate-

maatika-informaatika teaduskonna tudengitele suunatud aines kui ka teistele huvilistele suunatud aines "Programeerimise alused", mida 2013/2014 aasta kevadel kuulas 150 inimest. Autor loodab, et nende seas on ka mõned psühholoogia tudengid.

2. Pythonil on NumPy ja SciPy teekide näol väga hea arvutuslik tugi soovitud hinnangute implementeerimiseks.
3. Pythoni teeki on väga lihte integreerida mõne veebiraamistikuga, mis võiks olla loodava teegi üks rakenduskohti.
4. Python on vabavaraline ning tasuta kasutatav.
5. Sarnane programmiteek on juba implementeeritud keeles R [16].

Meetodid, mis leiavad λ_1 , λ_2 , λ_3 ja λ_4 , on implementeeritud nii programeerimiskeeles Python 2 kui ka programeerimiskeeles Python 3. Klass, mille abil saab leida *glb*, on implementeeritud sõltuvate teekide piirangute tõttu ainult programeerimiskeeles Python 2. Programmiteegi loomisel on kasutatud teeki SciPy [12] ja CVXPY [6]. Programmiteek on kättesaadav aadressil <https://github.com/vootelerotov/CTT-reliability-estimates>.

Autor märgib, et programmi, mis leiab hinnangu λ_4 , ajaline keerukus on $O(2^n)$, mis tähendab, et suurte küsimuste hulga korral ei ole selle kasutamine võimalik. Küll on aga sellisel juhul võimalik kasutada programme, mis leiavad lähendi hinnangule λ_4 (vaata [2]). Ülejäänud programmide keerukus on polünoomiaalne.

Lisas C on ära toodud hinnangute implementatsioonid programmeerimiskeeles Python 2.

5 Intervallskaal alternatiivne tõlgendamine

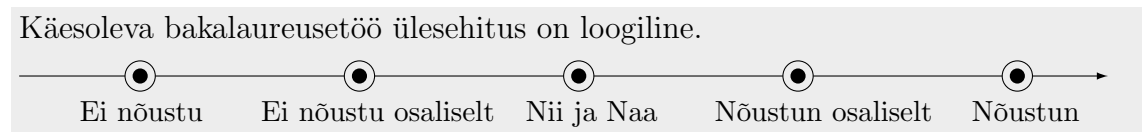
Antud töö alguses püstitasime küsimuse, kas leidub paremaid reliaabluse hinnanguid kui λ_3 .

Läheneme nüüd reliaabluse hinnangu parandamisele teisest suunast ning vaatleme, millist mõju avaldab hinnangule λ_3 vastuste teistsugune tõlgendamine intervallskaalal.

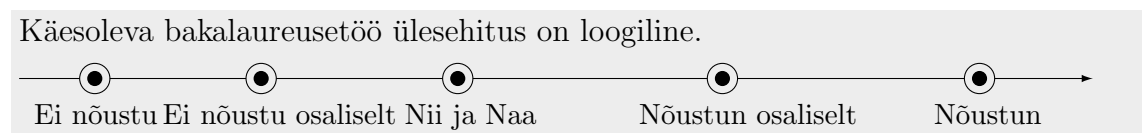
5.1 Vastuste tõlgendamine intervallskaalal

Käesoleva töö sissejuhatuses kirjeldasime viisi, kuidas vaikimisi teisendatakse vastused Likerti skaalalt intervallskaalale. Märkime, et selline tõlgendus eeldab, et vastused asuvad teineteisest sama „kaugel”. Arvame, et paljudel juhtudel ei pea see eeldus paika.

Sellest tähelepanekust lähtub järgmine küsimus – kas on võimalik leida paremat vastuste paigutust intervallskaalal kui vaikimisi meetod? Vaatleme kahte näidet:



Joonis 10: Näide sellest, kuidas paigutuvad hinnangud skaalal vaikimisi meetodit kasutades



Joonis 11: Näide hinnangute alternatiivsest paiknemisest skaalal

Vaikimisi tõlgendusele alternatiive välja pakkudes peab meil olema mingi karakteristik, mille alusel tõlgendust hinnata. Esitame järgmise idee: üritame leida hinnangute paiknemist skaalal nii, et küsimustiku reliaablus oleks võimalikult suur. Seame endale loomuliku piirangu: vastuste esialgne järjestus ei tohi muutuda.

5.2 Reliaabluse hindamine

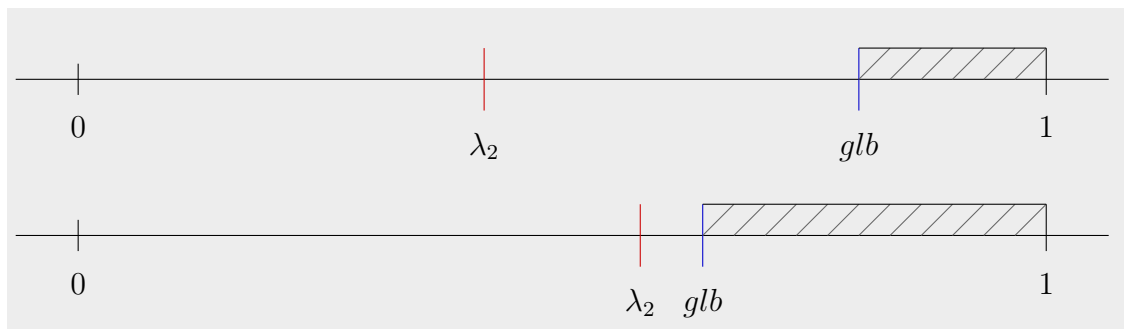
Kuna küsimustiku reliaabluse leidmine ei ole praktiliselt võimalik, siis kasutame selle lähendusena eelmises peatükis käsitletud alumisi tõkkeid. On selge, et parima tulemuse annab võimalikult täpse alumise tõkke kasutamine. Kuna sobiva tõlgenduse leidmine taandub optimeerimisülesandele, siis kasutame siinkohal konservatiivsemat alumist tõket λ_3 : selle arvutamise lihtsus, mis ei omanud fikseeritud tõlgenduse korral tänapäevase arvutusvõimuse juures erilist tähtsust, osutub siin oluliseks eeliseks. Lisaks on λ_3 psühholoogide seas laialt tuntud ning seda kasutava meetodi juurutamine praktikas oleks ilmselt lihtsam, kui mõne nende jaoks tundmatut hinnangut kasutava meetodi juurutamine.

5.3 Lähenemise korrektsus

Lugejal võib siinkohal tekkida kaks küsimust lähenemise korrektuse kohta.

Esiteks, kas selline lähenemine tagab meile küsimustiku kõrgema reliaabluse? Vastus sellele küsimusele on kahjuks eitav – tõepoolest, on võimalik, et me saame teistsuguse tõlgendusega küsimustiku, mille reliaabluse alumine tõke on suurem kui vaikimisi tõlgendusega küsimustiku puhul, kuid mille reliaablus on samal ajal madalam. Kuna meil aga ei ole võimalik reliaablust täpselt mõõta, siis ei ole meil ka otsest põhjust uskuda, et ühe tõlgendusega saadud reliaablus on parem kui teisega.

Küll aga kerkib siinkohal esile praktilisem probleem seoses λ_3 kasutamisega reliaabluse hindamiseks. Nimelt võib leida aset olukord, kus meie pakutava tõlgenduse korral on λ_3 kõrgem kui vaikimisi tõlgenduse korral, kuid mõni tihedam alumine tõke on madalam. Illustreerime seda olukorda näitega.



Joonis 12: Olukord, kus alternatiivne tõlgendus suurendab alumist tõket λ_3 ja vähendab alumist tõket glb .

Joonisel kujutatud olukord on kindlasti ebasoovitav. Lihtsam võimalik lahendus on uue tõlgenduse korral leida teised alumised tõkked ning juhul, kui kõige tihedam alumine tõke on vaikimisi tõlgenduse korral suurem kui leitud tõlgenduse korral, siis jääda vaikimisi tõlgenduse juurde. Keerukam lahendus oleks kasutada optimeerimisel mõnda tihedamat alumist tõket – selle lähenemise puuduseks on aga vajadus lahendada keerukam optimeerimisülesanne.

Teiseks, kuidas mõjutab vastuste tavapärasest erinev tõlgendus küsimustiku valideerimist? Selle küsimuse korral märgime ära, et kuigi vaikimisi definitsioon on laialdaselt levinud, on ta siiski täiesti arbitraarne ning seega samaväärne meie poolt pakutud tõlgendusega – meil ei ole mingit kindlat põhjust uskuda, et ühe tõlgenduse korral on valideerimine parem kui teise korral. Ühelt poolt on selline vastus mittera-

huldav – meie pakutud lahendus ei pruugi olla parem kui olemasolev. Teisalt ei pruugi vaikimisi kasutatav tõlgendus olla parem kui meie poolt pakutav.

Meie pakutava tõlgenduse väikese eelisena võib siiski ära tuua järgneva. Kuna enamiku valiidsuse definitsioonide korral on valiidsus ülevalt tõkestatud mingi reliaablusest sõltuva funktsiooniga (kõige levinuma definitsiooni korral on tõkkeks reliaabluse positiivne ruutjuur), siis reliaabluse alumise tõkke abil saame midagi öelda valiidsuse kõige väiksema võimaliku ülemise tõkke kohta. Paneme tähele, et kuna teame ainult reliaabluse alumist tõket ja mitte reliaablust ennast, siis ei ole meil mingit põhjust väita, et vähim võimalik ülemine tõke on ka küsimustiku valiidsuse ülemiseks tõkkeks.

5.4 Alternatiivse tõlgenduse matemaatiline sõnastus

Olgu meil populatsioon I ning k küsimusega küsimustik, kus iga küsimuse kohta palutakse hinnangut l -palli Likerti skaalal. Nummerdama küsimused naturaalarvudega 1 kuni k . Olgu meil juhuslikud suurused K_1, K_2, \dots, K_k , mille määramispiirkonnaks on kõik erinevad võimalikud vastuste kombinatsioonid küsimustikule ning muutumispiirkonnaks on hulk $\{1, 2, \dots, l\}$ ning tõlgendame juhusliku suurust kui küsimusele k antud vastuse tõlgendust vaikimisi kasutataval intervallskaalal. Toome sisse ka tähistused $p_{i\alpha}, i \in \{1, 2, \dots, n\}, \alpha \in \{1, 2, \dots, l\}$, kus $p_{i\alpha}$ tähistab põhjal tõenäosusust, et juhusliku suuruse K_i väärtus on α .

Tuletame meelde λ_3 definitsiooni üle vektori (K_1, K_2, \dots, K_k) kovariatsioonimaatriksi Σ_K :

$$\lambda_3 = \left(\frac{k}{k-1}\right) \left(1 - \frac{\sum_{i=1}^k \Sigma_{Kii}}{\sum_{i=1}^k \sum_{j=1}^k \Sigma_{Kij}}\right)$$

Võttes arvesse, et $\Sigma_{Kij} = \text{cov}(K_i, K_j)$, ning definitsiooni järgi $\text{cov}(X, X) = \epsilon[(X - \epsilon[X])(X - \epsilon[X])]$, saame, et

$$\lambda_3 = \left(\frac{k}{k-1}\right) \left(1 - \frac{\sum_{i=1}^k \Sigma_{Kii}}{\sum_{i=1}^k \sum_{j=1}^k \Sigma_{Kij}}\right) = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k D(K_i)}{\sum_{i=1}^k \sum_{j=1}^k \text{cov}(K_i, K_j)}\right)$$

Soovime leida sellise viisi vastuste tõlgendamiseks, et λ_3 oleks maksimaalne. Kuju-tame juhuslikud suurused K_1, K_2, \dots, K_n juhuslikeks suurusteks M_1, M_2, \dots, M_n , kusjuures juhusliku suuruse M_i määramispiirkond ühtib suuruse K_i määramis-piirkonnaga ning M_i saab väärtusi hulgast $M_i = \{\mu_{1i}, \mu_{2i}, \dots, \mu_{li}\}$ $i \in 1, 2, \dots, n$. Lisaks kehtigu järgnevad kitsendused:

$$\mu_{i1} < \mu_{i2} < \dots < \mu_{il},$$

$$K_i(\omega) = j \implies M_i(\omega) = \mu_{ij}, \forall j, 1 \leq j \leq l.$$

Siinkohal märgime, et hinnangute tõlgenduste juures ei huvita meid absoluutne, vaid suhteline paigutus. Seda arvesse võttes lisame järgnevad normaliseerivad kit-sendused:

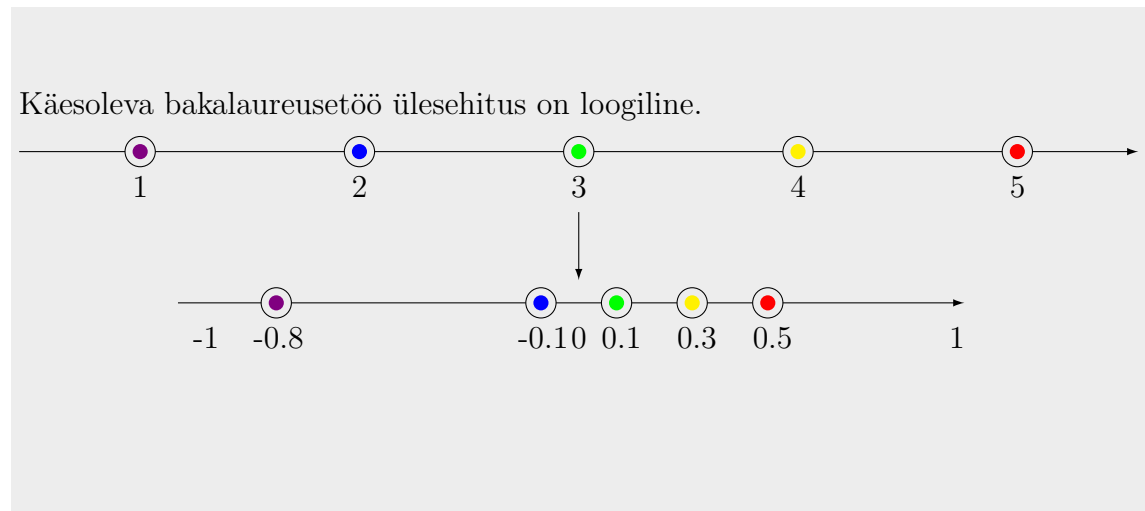
$$E(L_i) = \sum_{j=1}^l p_{ij} \mu_{ij} = 0, \quad (5.1)$$

$$D(L_i) = \sum_{j=1}^l p_{ij} (\mu_{ij} - E(L_i))^2 = \sum_{j=1}^l p_{ij} \mu_{ij}^2 = 1. \quad (5.2)$$

Paneme tähele, et selliste piirangute korral esitub λ_3 lihtsamal kujul. Tõepoo-lest,

$$\lambda_3 = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k D(M_i)}{\sum_{i=1}^k \sum_{j=1}^k \text{cov}(M_i, M_j)} \right) = \frac{k}{k-1} \left(1 - \frac{k}{\sum_{i=1}^k \sum_{j=1}^k \text{cov}(M_i, M_j)} \right).$$

Siit näeme, et sellisel juhul taandub λ_3 maksimeerimine avaldise $\sum_{i=0}^n \sum_{j=0}^n \text{COV}(M_i, M_j)$ maksimeerimisele.



Joonis 13: Näide hulkade $\text{ran}(K_i)$ ja $\text{ran}(M_i)$ suhestumisest

Olgu meil tõenäosuste maatriks P , kusjuures

$$P = \begin{pmatrix} p_{(11)(11)} & p_{(11)(12)} & \cdots & p_{(11)(1l)} & p_{(11)(21)} & p_{(11)(22)} & \cdots & p_{(11)(kl)} \\ p_{(12)(11)} & p_{(12)(12)} & \cdots & p_{(12)(1l)} & p_{(12)(21)} & p_{(12)(22)} & \cdots & p_{(12)(kl)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{(kl)(11)} & p_{(kl)(12)} & \cdots & p_{(kl)(1l)} & p_{(kl)(21)} & p_{(kl)(22)} & \cdots & p_{(kl)(kl)} \end{pmatrix},$$

kus $p_{(i\alpha)(j\beta)}$, $i, j \in \{1, 2, \dots, k\}$, $\alpha, \beta \in \{1, 2, \dots, l\}$ tähistab tõenäosust, et juhuslik suurus K_i saab väärtuse α ja juhuslik suurus K_j saab väärtuse β .

Defineerime vektori x , nii et

$$x = (\mu_{11}, \mu_{12}, \dots, \mu_{1l}, \mu_{21}, \mu_{22}, \dots, \mu_{kl}).$$

Siis $xPx^T = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(M_i, M_j)$. Veendume selles:

$$\begin{aligned}
xPx^T &= (\mu_{11} \quad \mu_{12} \quad \cdots \quad \mu_{kl}) \begin{pmatrix} p_{(11)(11)} & p_{(11)(12)} & \cdots & p_{(11)(kl)} \\ p_{(12)(11)} & p_{(12)(12)} & \cdots & p_{(12)(kl)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{(kl)(11)} & p_{(kl)(12)} & \cdots & p_{(kl)(kl)} \end{pmatrix} \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{kl} \end{pmatrix} = \\
&= \begin{pmatrix} \sum_{j=1}^n \sum_{r=1}^l \lambda_{jr} p_{(jr)(11)} & \sum_{j=1}^n \sum_{r=1}^l \lambda_{jr} p_{(jr)(12)} & \cdots & \sum_{j=1}^n \sum_{r=1}^l \lambda_{jr} p_{(jr)(kl)} \end{pmatrix} \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{kl} \end{pmatrix} = \\
&= \sum_{j=1}^n \sum_{r=1}^l \lambda_{jr} p_{(jr)(11)} + \sum_{j=1}^n \sum_{r=1}^l \lambda_{jr} p_{(jr)(12)} + \cdots + \sum_{j=1}^n \sum_{r=1}^l \lambda_{jr} p_{(jr)(kl)} = \\
&= \sum_{i=1}^k \sum_{s=1}^l \sum_{j=1}^k \sum_{r=1}^l \lambda_{jr} p_{(jr)(is)} \lambda_{is} = \sum_{i=1}^k \sum_{j=1}^k \sum_{s=1}^l \sum_{r=1}^l \lambda_{jr} p_{(jr)(is)} \lambda_{is} = \sum_{i=1}^k \sum_{j=1}^k E(M_i M_j) = \\
&\stackrel{(5.1)}{=} \sum_{i=1}^k \sum_{j=1}^k E(M_i M_j) - E(M_i) E(M_j) = \sum_{i=1}^k \sum_{j=1}^k cov(M_i, M_j)
\end{aligned}$$

Eelneva põhjal piisab meile, arvestades eelnevalt äratoodud piiranguid L_i keskväärtusele ja dispersioonile ning pannes tähele, et mõlemat piirangut on võimalik lihtsasti kirja panna maatriksite korrutisena, suurima võimaliku λ_3 leidmiseks ruutvõrrandi xPx^T maksimiseerimisest. Seega saame püstitada järgneva optimeerimisprobleemi, mille lahendus annab meile otsitavad tõlgendused.

$$\begin{aligned}
&\min x^T(-P)x \\
&R_i^T x = 0, i \in 1, 2, \dots, k \\
R_i &= (\underbrace{0, 0, \dots, 0, 0}_{(i-1)*l}, p_{i1}, p_{i2}, p_{i3}, p_{i4}, \dots, p_{il}, \underbrace{0, 0, \dots, 0, 0}_{(k-i)*l}) \quad (5.3) \\
x^T Q_i x &= 1, i \in \{1, 2, \dots, k\}, \quad Q_i = \begin{pmatrix} q_{11} & 0 & \cdots & 0 \\ 0 & q_{12} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_{kl} \end{pmatrix} \\
q_{jk} &= \begin{cases} 0 & j \neq i \\ p_{ik} & i = j \end{cases}
\end{aligned}$$

5.5 Hinnang saadud tõlgendusele

Selles peatükis välja pakutud idee kasutamise esimeseks eelduseks on püstitatud optimeerimisülesande lahendamine. Antud ülesande puhul on tegemist ruutkit-sendustega ruutprogrameerimisülesandega, mille lahendamine on üldjuhul raske^⑤. Lisaks tundub, et antud ülesanne ei kuulu mõnda lihtsamasse alamülesannete klas-si.

Küll on sellist tüüpi ülesandeid võimalik lõdvendada, levinud lõdvenduseks on probleemi teisendamine mittenegatiivseks optimeerimisülesandeks [1, 2] – see konkreetne lõdvendus osutus antud probleemi puhul aga liiga lõdvaks.

Autor leiab, et juhul, kui õnnestub leida efektiivne viis optimeerimisülesande lähendamiseks, väärib tutvustatud idee edasist uurimist – λ_3 tuntus tähendab seda, et juhul, kui on põhjust uskuda, et optimeerimisülesande lahendamisel saadud tõlgendused suurendavad testi reliaablust, võib loota, et väljapakutud meetod leiab ka praktikas kasutamist.

^⑤ Täpsemalt on tegemist *NP-Hard* probleemiga: kui sellele ülesandele leidub polünomiaalse keerukusega lahendusalgorithm, siis leidub polünomiaalse keerukusega algoritm ka paljudele teiste rasketele probleemidele

Kokkuvõte

Antud bakalaureusetöös üritasime lahendada praktilist probleemi – kas reliaabluse hindamiseks leidub paremat hinnangut, kui hetkel praktikas kasutatakse λ_3 , tuntud ka kui Cronbachi α ?

Töö esimeses osas tutvustasime lugejale küsimustikega seotud terminoloogiat, sealhulgas valiidsuse ja reliaabluse mõisteid ning esitasime vajaliku taustinfo probleemi mõtestamiseks. Teises peatükis jõudsime reliaabluse matemaatilise definitsioonini ning tutvustasime lugejale ühtlasi klassikalise testiteooria põhitulemus. Lisaks vaatlesime, kuidas seostuvad küsimustiku vaadeldavate tulemuste, tegelike tulemuste ja mõõtevigade kovariatsioonimaatriksid küsimustiku reliaablusega. Kolmandas osas tuletasime viis reliaabluse alumist tõket: λ_1 , λ_2 , λ_3 , λ_4 ja *glb*. Järgnevalt kirjeldasime töö käigus loodud programmeerimise ning põhjendasime selle implementeerimisel tehtud valikuid. Töö viimases osas vaatlesime ühte võimalikku viisi küsimustiku reliaabluse parandamiseks, leides vastuste sobivama paigutuse intervallskaalal.

Autor loodab, et töö pakkus midagi nii psühholoogidele kui ka matemaatikutele. Psühholoogidel võib abi olla töös tutvustatud reliaabluse hinnangutest ning nende hinnangute praktikas leidmiseks loodud programmeerimisest. Matemaatikute jaoks pakkus loodetavasti huvi nii töös tutvustatud klassikaline testiteooria, mis annab matemaatiliselt ilusa raamistiku nii küsimustike analüüsimiseks kui ka töö viimases osas esitatud idee küsimustike reliaabluse suurendamiseks.

Viited

- [1] ASPREMONT, A.; BOYD, S. Relaxations and Randomized Methods for Nonconvex QCQPs. <http://www.stanford.edu/class/ee392o/relaxations.pdf>, 2003.
- [2] BENTON, T. An empirical assessment of Guttman ' s Lambda 4 reliability coefficient. <http://www.cambridgeassessment.org.uk/Images/141299-an-empirical-assessment-of-guttman-s-lambda-4-reliability-coefficient.pdf>, 2013.
- [3] CRONBACH, L. J. Test "Reliability": Its Meaning and Determination. *Psychometrika* 12, 1 (1947), 1–16.
- [4] CRONBACH, L. J. Coefficient Alpha and the Internal Structure of Tests. *Psychometrika* 16, 3 (1951), 297–330.
- [5] CRONBACH, L. J.; SHAVELSON, R. J. My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement* 64, 3 (June 2004), 391–418.
- [6] DIAMOND, S., CHU, E.; BOYD, S. CVXPY: A Python-Embedded Modeling Language for Convex Optimization, 2014.
- [7] EDMONDSON, D. R. Likert Scales: A History. In *Conference on Historical Analysis & Research in Marketing Proceedings* (2005).
- [8] GUTTMAN, L. A Basis for Analyzing Test-Retest Reliability. *Psychometrika* 10, 4 (1945), 255–282.
- [9] HENSON, R. K. Understanding Internal Consistency Reliability Estimates: A Conceptual Primer on Coefficient Alpha. *Measurement & Evaluation in Counseling & Development* 34, 3 (2001), 177.
- [10] JACKSON, P. H.; AGUNWAMBA, C. C. Lower Bounds for the Reliability of the Total Score on a Test Composed of Non-Homogeneous Items: I: Algebraic Lower Bounds. *Psychometrika* 42, 2 (1977), 567–578.
- [11] JAMIESON, S. Likert scales: how to (ab)use them. *Medical education* 38, 12 (Dec. 2004), 1217–8.
- [12] JONES, E., OLIPHANT, T.; PETERSON, P. SciPy: Open Source Scientific Tools for Python.
- [13] LORD, F.; NOVICK, M. *Statistical theories of mental test scores*. 1968.
- [14] NOVICK, M. R. The axioms and principal results of classical test theory. *Journal of Mathematical Psychology* 3, 1 (Feb. 1966), 1–18.

- [15] PÄRNA, K. *Tõenäosusteooria algkursus*. Tartu Ülikooli kirjastus, Tartu, 2013.
- [16] REVELLE, W. *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois, 2014.
- [17] REVELLE, W.; ZINBARG, R. E. Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijsma. *Psychometrika* 74, 1 (Dec. 2008), 145–154.
- [18] SIJTSMA, K. On the use, the misuse, and the very limited usefulness of cronbach’s alpha. *Psychometrika* 74, 1 (2009), 107–120.
- [19] SIJTSMA, K. Reliability Beyond Theory and Into Practice. *Psychometrika* 74, 1 (2009), 169–173.
- [20] SMITH, N. Example of Likert Scale. http://en.wikipedia.org/wiki/File:Example_Likert_Scale.svg.
- [21] TAMMERAID, I. *Lineaaralgebra rakendused*. TTÜ Kirjastus, Tallinn, 1999.
- [22] TEN BERGE, J. M.; SOCAN, G. The Greatest Lower Bound to the Reliability of a Test and the Hypothesis of Unidimensionality. *Psychometrika* 69, 4 (2004), 613–625.
- [23] UT. Õisi tagasiside. <http://is.ut.ee>.
- [24] ZIMMERMAN, D. W. Probability spaces, hilbert spaces, and the axioms of test theory. *Psychometrika* 40, 3 (1975), 395–412.
- [25] ZIMMERMAN, D. W. Test Theory With Minimal Assumptions. *Educational and Psychological Measurement* 36, 1 (Apr. 1976), 85–96.

Lisad

A Näited küsimustikust Likerti skaalal

3. Õppejõu hoiak õpetamisel oli õppimist toetav ja üliõpilaste suhtes avatud.

☐ nõustun igati ☐ pigem nõustun ☐ pigem ei nõustu ☐ ei nõustu üldse ☐ nii ja naa

3.1. Palun kommenteerige ja/või põhjendage oma hinnangut.

4. Õppejõud õpetas õppeainet meisterlikult (huvi äratamine, esitatu arusaadavus, kaasahaaravus jms).

☐ nõustun igati ☐ pigem nõustun ☐ pigem ei nõustu ☐ ei nõustu üldse ☐ nii ja naa

4.1. Palun kommenteerige ja/või põhjendage oma hinnangut.

5. Õppejõu antud või soovitatud õppematerjalid olid sisu, vormistuse ja sobivuse poolest asjakohased.

☐ nõustun igati ☐ pigem nõustun ☐ pigem ei nõustu ☐ ei nõustu üldse ☐ nii ja naa

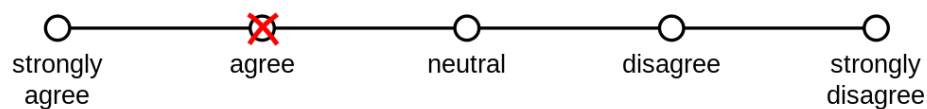
5.1. Palun kommenteerige ja/või põhjendage oma hinnangut.

6. Õppejõud andis piisavalt tagasisidet minu töö tulemuste kohta aine läbimisel.

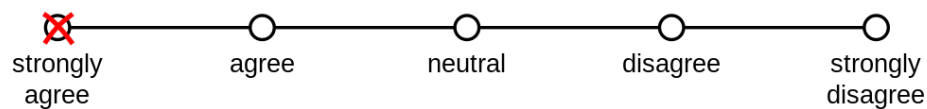
☐ nõustun igati ☐ pigem nõustun ☐ pigem ei nõustu ☐ ei nõustu üldse ☐ nii ja naa

Joonis 14: Näide Tartu Ülikooli õppeinfo süsteemi tagasiside ankeedist, kus rakendatakse Likerti skaalat [23]

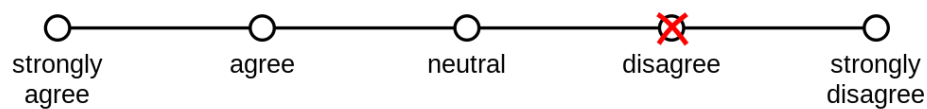
1. Wikipedia has a user friendly interface.



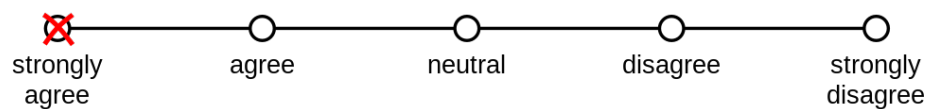
2. Wikipedia is usually my first resource for research.



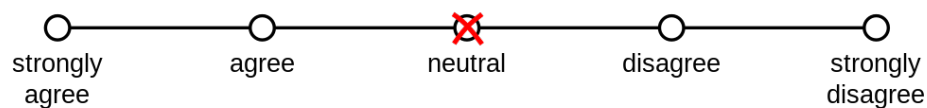
3. Wikipedia pages generally have good images.



4. Wikipedia allows users to upload pictures easily.



5. Wikipedia has a pleasing color scheme.



Joonis 15: Näide küsimustikust, kus on rakendatud Likerti skaalat; küsimused on paigutatud nende järjestikulisuse rõhutamiseks teljele [20]

B Empiirilised näited

Järgnevad kovariatsioonimaatriksid on leitud ühe Tartu ülikooli psühholoogide poolt läbiviidud küsimustiku andmete põhjal, jagadas küsimustiku alamküsimustikeks.

Kovariatsioonimaatriks											
3.00	1.52	1.73	1.21	0.69	1.40	-0.53	-0.64	-0.75	-0.56	-0.57	-0.83
1.52	3.25	1.37	0.70	0.78	1.01	-0.31	-0.21	-0.17	-0.39	-0.29	-0.65
1.73	1.37	3.25	1.43	0.85	1.62	-0.71	-0.77	-0.94	-0.98	-0.65	-1.20
1.21	0.70	1.43	2.79	0.16	1.00	-1.03	-1.23	-0.98	-0.87	-0.82	-0.85
0.69	0.78	0.85	0.16	2.67	0.82	-0.11	0.20	-0.10	0.10	0.40	-0.09
1.40	1.01	1.62	1.00	0.82	2.37	-0.52	-0.48	-0.84	-0.55	-0.37	-0.73
-0.53	-0.31	-0.71	-1.03	-0.11	-0.52	1.59	1.25	0.52	0.87	0.65	0.93
-0.64	-0.21	-0.77	-1.23	0.20	-0.48	1.25	2.38	0.77	1.12	0.92	1.02
-0.75	-0.17	-0.94	-0.98	-0.10	-0.84	0.52	0.77	2.30	0.81	0.75	0.68
-0.56	-0.39	-0.98	-0.87	0.10	-0.55	0.87	1.12	0.81	2.14	1.00	1.25
-0.57	-0.29	-0.65	-0.82	0.40	-0.37	0.65	0.92	0.75	1.00	2.09	0.69
-0.83	-0.65	-1.20	-0.85	-0.09	-0.73	0.93	1.02	0.68	1.25	0.69	2.03

Hinnang	λ_1	λ_2	λ_3	λ_4	glb
Väärtus	0.39	0.60	0.43	0.73	0.77

Joonis 16: Kahe erineva kuue küsimusega alamküsimustiku kombineerimisel saadud kovariatsioonimaatriks koos reliaabluse hinnangutega

Kovariatsioonimaatriks												
Variable	1	2	3	4	5	6	7	8	9	10	11	12
	3.00	1.52	1.73	1.21	0.69	1.40	1.39	0.90	1.27	1.15	0.58	0.83
	1.52	3.25	1.37	0.70	0.78	1.01	1.10	1.73	0.94	0.85	0.56	0.70
	1.73	1.37	3.25	1.43	0.85	1.62	1.19	0.61	1.85	1.28	0.61	0.85
	1.21	0.70	1.43	2.79	0.16	1.00	0.64	0.26	0.93	1.49	0.14	0.53
	0.69	0.78	0.85	0.16	2.67	0.82	0.77	0.99	0.64	0.10	1.61	0.56
	1.40	1.01	1.62	1.00	0.82	2.37	0.98	0.74	1.17	1.02	0.74	0.90
	1.39	1.10	1.19	0.64	0.77	0.98	2.39	1.26	1.12	0.66	0.61	0.72
	0.90	1.73	0.61	0.26	0.99	0.74	1.26	2.91	0.55	0.43	0.82	0.77
	1.27	0.94	1.85	0.93	0.64	1.17	1.12	0.55	2.52	0.81	0.45	0.80
	1.15	0.85	1.28	1.49	0.10	1.02	0.66	0.43	0.81	2.19	0.25	0.55
	0.58	0.56	0.61	0.14	1.61	0.74	0.61	0.82	0.45	0.25	2.43	0.54
	0.83	0.70	0.85	0.53	0.56	0.90	0.72	0.77	0.80	0.55	0.54	1.35

Hinnang	λ_1	λ_2	λ_3	λ_4	<i>glb</i>
Väärtus	0.79	0.87	0.86	0.93	0.93

Joonis 17: Kahe sama konstruktsiooni mõõtvat kuue küsimusega alamküsimustike kombineerimisel saadud kovariatsioonimaatriks koos reliaabluse hinnangutega

Kovariatsioonimaatriks											
1.99	0.39	0.37	0.98	0.85	0.70	-0.48	-0.49	-0.39	-0.43	-0.25	-0.27
0.39	2.15	0.38	0.79	0.50	0.28	-0.60	-0.39	-0.27	-0.05	-0.19	-0.05
0.37	0.38	2.12	0.45	0.74	0.70	-0.41	-0.10	-0.06	0.06	-0.02	0.07
0.98	0.79	0.45	3.06	0.95	0.34	-0.59	-0.64	-0.36	-0.14	-0.04	0.05
0.85	0.50	0.74	0.95	2.90	0.77	-0.32	-0.58	-0.05	0.05	0.22	0.32
0.70	0.28	0.70	0.34	0.77	2.65	-0.47	-0.45	-0.45	-0.37	-0.39	-0.38
-0.48	-0.60	-0.41	-0.59	-0.32	-0.47	1.99	0.58	0.51	0.38	0.31	0.47
-0.49	-0.39	-0.10	-0.64	-0.58	-0.45	0.58	1.86	0.59	0.30	0.03	0.31
-0.39	-0.27	-0.06	-0.36	-0.05	-0.45	0.51	0.59	1.63	0.63	0.45	0.81
-0.43	-0.05	0.06	-0.14	0.05	-0.37	0.38	0.30	0.63	1.41	0.63	1.01
-0.25	-0.19	-0.02	-0.04	0.22	-0.39	0.31	0.03	0.45	0.63	1.79	1.11
-0.27	-0.05	0.07	0.05	0.32	-0.38	0.47	0.31	0.81	1.01	1.11	2.33

Hinnang	λ_1	λ_2	λ_3	λ_4	<i>glb</i>
Väärtus	0.39	0.53	0.43	0.68	0.70

Joonis 18: Kahe erineva kuue küsimusega alamküsimustiku kombineerimisel saadud kovariatsioonimaatriks koos reliaabluse hinnangutega

C Programmiteegi implementatsioon

```
1 def calculate_lambda_1(cov_matrix):
2     u"""
3     Based on Guttman,1945, for calculations using Jackson,
4     Aguvwamba,1977,[569].
5     """
6     n = check_that_matrix_is_square_and_fix_n(cov_matrix)
7     def sum_of_main_diagonal():
8         summ = 0
9         for i in xrange(n):
10             summ += cov_matrix[i,i]
11         return summ
12
13     return 1 - (sum_of_main_diagonal()/
14                 calculate_sum_of_covariance_matrix_elements(cov_matrix, n))
```

Programm 1: Funktsioon hinnangu λ_1 leidmiseks

```
1 def calculate_lambda_2(cov_matrix):
2     u"""
3     Using Guttman,1945,[259] definition.
4     """
5     def sum_of_squares_of_non_diagonal_matrix_elements():
6         summ = 0
7         for i in xrange(n):
8             for j in xrange(i):
9                 summ += cov_matrix[i,j]**2
10        return 2*summ
11
12    n = len(cov_matrix)
13    l1 = calculate_lambda_1(cov_matrix)
14    sum_of_squares = sum_of_squares_of_non_diagonal_matrix_elements()
15    under_square_root = ((n)/(n-1)* sum_of_squares)**(1/2)
16    return l1 + (under_square_root/
17                calculate_sum_of_covariance_matrix_elements(cov_matrix))
```

Programm 2: Funktsioon hinnangu λ_2 leidmiseks

```
1 def calculate_lambda_3(cov_matrix):
2     number_of_items = len(cov_matrix)
3     return (number_of_items/(number_of_items-1))*calculate_lambda_1(cov_matrix)
```

Programm 3: Funktsioon hinnangu λ_3 leidmiseks

```

1 import numpy as np
2 def calculate_lambda_4(cov_matrix):
3     u"""
4     Lambda 4, based on Guttman, 1945, using the implementation of Jackson,
5     Agunwamba, 1977.
6     Based on idea that we try to best possible split (u is a vector,
7     with elements either 1 or -1, ones will be in one split,
8     minus ones in another) .
9     NAIVE IMPLEMENTATION, works with relatively low N,
10    better approaches available, see Benton, 2013.
11    """
12
13    def objective_function(u):
14        return u.T.dot(cov_matrix.dot(u))
15
16    def try_vectors():
17        u"""
18        Idea : generate half of all possible vectors of length n,
19        such that if vector v is in, vector -v is not.
20        Using binary representation as string,
21        from 2**(n-1) to 2**n, coding "0" to "-1"
22        """
23
24        smallest = np.Infinity
25        result_vector = []
26        l = []
27        for i in xrange(2**(n-1), 2**n):
28            binary_of_i = np.binary_repr(i, width=n)
29            binary_of_i_int = [ 1 if x == "1" else -1 for x in binary_of_i]
30            u = np.array(binary_of_i_int)
31            result = objective_function(u)
32            if result < smallest:
33                smallest = result
34                result_vector = u
35        return smallest
36
37    def calc_lambda(smallest):
38        return 1 - (smallest/
39                    util.calculate_sum_of_covariance_matrix_elements(cov_matrix, n))
40
41    n = util.check_that_matrix_is_square_and_fix_n(cov_matrix)
42    return calc_lambda(try_vectors())

```

Programm 4: Funktsioon hinnangu λ_4 leidmiseks

```

1 import cvxpy as cv
2 import numpy as np

3 u"""
4 Idea: Observed covariance matrix is of observed scores.
5 That matrix is sum of true score covariance matrix and error covariance matrix.
6 As errors don't correlate with anything else than themselves,
7 the last is diagonal matrix. So,  $C_{obs} = C_{true} + C_{error}$ 
8 We look to maximize error (to find the worst possible case),
9 that is the same as minimizing the trace of  $C_{true}$ .
10 More can be found in Jackson, Agunwamba, 1977 and Woodhouse, Jackson, 1977.
11 """

12 class Glb(object):
13     def __init__(self, cov_matrix):
14         self._cov_matrix = cov_matrix
15         self._constraints = []
16         self._n = util.check_that_matrix_is_square_and_fix_n(cov_matrix)
17         self._cov_matrix = np.matrix(cov_matrix)
18         self._generate_matrix_variables()
19         self._divide_covariance_matrix_of_into_error_and_true_score()
20         self._fix_error_cov_matrix_diagonal_fields_constraints()
21         self._error_cov_matrix_non_diagonal_fields_are_0()
22         self._prepare_program()
23         self._solve_program()
24         self._answer = self._calculate_Glb()

25     def _generate_matrix_variables(self):
26         self._true_cov_matrix = cv.semidefinite(self._n,
27             name=u"true_cov_matrix")
28         self._error_cov_matrix = cv.Variable(self._n, self._n,
29             name=u"error_cov_matrix")

30     def _divide_covariance_matrix_of_into_error_and_true_score(self):
31         u"""
32         Constraint that  $C_{obs} = C_{true} + C_{error}$ 
33         """
34         self._constraints.append(self._error_cov_matrix+
35             self._true_cov_matrix == self._cov_matrix)

36     def _fix_error_cov_matrix_diagonal_fields_constraints(self):
37         u"""
38         Error covariance matrix must be semidefinite,
39         meaning that all elements on the diagonal must be positive.
40         """
41         for i in xrange(self._n):
42             self._constraints.append(self._error_cov_matrix[i,i] >= 0)

```

```

43 def _error_cov_matrix_non_diagonal_fields_are_0(self):
44     u"""
45     Covariance between different errors is 0
46     """
47     for i in xrange(self._n):
48         for j in xrange(self._n):
49             if (i != j):
50                 self._constraints.append(
51                     self._error_cov_matrix[i,j] == 0)

52 def _prepare_program(self):
53     u"""
54     Minimize trace of true score covariance matrix
55     """
56     self._p = cv.Problem(cv.Minimize(
57         sum([self._true_cov_matrix[i,i] for i in xrange(self._n)])),
58         self._constraints)
59     if not self._p.is_dcp():
60         raise Exception(u"Non-DCP glb, something went terrible wrong")

61 def _solve_program(self):
62     try:
63         self._p.solve()
64     except:
65         raise Exception(u"Program was unable to find glb")

66 def _calculate_Glb(self):
67     return 1 - (sum(
68         [self._error_cov_matrix.value[i,i] for i in xrange(self._n)])/
69         util.calculate_sum_of_covariance_matrix_elements(
70             self._cov_matrix,self._n))

71 def get_answer(self):
72     return self._answer

```

Programm 5: Klass hinnangu *glb* leidmiseks

```

def check_that_matrix_is_square_and_fix_n(input_matrix):
    def check_that_matrix_is_square():
        for row in input_matrix:
            if len(row) != n:

```

```

        raise Exception(u"Matrix not square or missing data")
    n = len(input_matrix)
    check_that_matrix_is_square()
    return n

def fix_number_of_rows(input_matrix):
    def check_that_matrix_is_complete():
        for row in input_matrix:
            if len(row) != n:
                raise Exception(u"Matrix not square or missing data")
    n = len(input_matrix[0])
    check_that_matrix_is_complete()
    return n

def calculate_sum_of_covariance_matrix_elements(cov_matrix, n = None):
    if n == None:
        n = check_that_matrix_is_square_and_fix_n(cov_matrix)
    summ = 0
    for i in xrange(n):
        for j in xrange(n):
            summ += cov_matrix[i,j]
    return summ

```

Programm 6: Eelnevates programmides kasutatud abifunktsioonid

D Litsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks.

Mina, Vootele Rõtov (sünnikuupäev: 11.11.1988)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
Valikvastustega küsimustike reliaabluse hindamine,
mille juhendaja on Margus Niitsoo,
 - (a) reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace'i lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - (b) üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartu, 3. juuni 2014. a.