

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 hinnangute tuletamiseks. Seejärel tuletatakse viis reliaabluse hinnangut, hinnatakse nende kasutatavust praktikas ja tutvustatakse autori loodud programmeerimisekeeli, 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, mis annavad paremaid 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 is to give an overview of different possibilities of estimating 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 from the administrated questionnaire. Finally, the author proposes an approach for improving the quantitative interpretation of a questionnaire.

The thesis gives an overview of the foundation of the Classical Test Theory in Estonian from mathematical viewpoint and introduces different possibilities for estimating the reliability of a questionnaire, some of which give better results 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 tõkked | 24 |
| 3.1 Alumine tõke λ_1 | 24 |
| 3.2 Alumine tõke λ_2 | 25 |
| 3.3 Alumine tõke λ_3 | 28 |
| 3.4 Alumine tõke λ_4 | 31 |
| 3.5 Alumine tõke glb | 31 |
| 3.6 Hinnangute kasutatavus praktikas | 32 |
| 4 Programmiteek | 33 |
| 5 Intervallskaal alternatiivne tõlgendamine | 34 |
| 5.1 Vastuste tõlgendamine intervallskaalal | 34 |
| 5.2 Reliaabluse hindamine | 35 |
| 5.3 Lähenemise korrektsus | 35 |
| 5.4 Alternatiivse tõlgenduse matemaatiline sõnastus | 36 |
| 5.5 Hinnang saadud tõlgendusele | 40 |
| Kokkuvõte | 41 |
| Viited | 42 |
| Lisa | 44 |
| Lisa A Näited küsimustikust Likerti skaalal | 44 |

| | | |
|--------|---|----|
| Lisa B | Empiirilised näited | 46 |
| Lisa C | Programmiteegi implementatsioon | 47 |
| Lisa D | Litsents | 52 |

Sissejuhatus

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

Sellest lähtuvalt on töö peamiseks eesmärgiks pakkuda 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 bakalau-reuseõppe läbinud matemaatikatudengile – meeldetuletuseks on valdkonna spetsiifilised matemaatilised definitsioonid ja tulemused töö käigus jooksvalt ära toodud.

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

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 [5], 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 13 ja 14

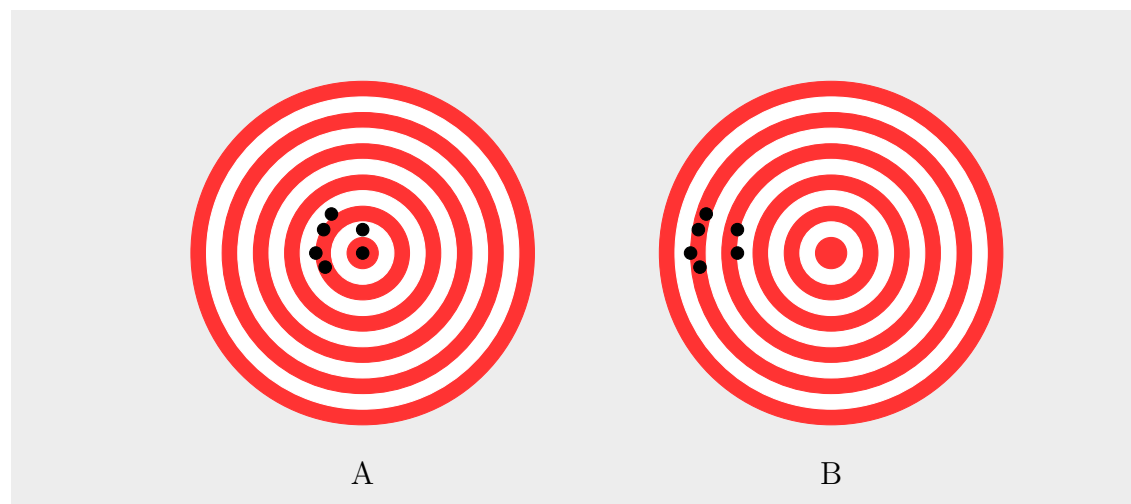
Lugejal võib tekkida õigustatud küsimus, kuidas põhjendab autor Likerti skaala käsitlemist intervallskaalana, kui Likerti skaala on olemuselt järjestikaskaala ning selle tõlgendamine intervallskaalana on vastuoluline küsimus, näiteks [9]. Siinkohal tõdeme, et Likerti skaala tõlgendamine intervallskaalana 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 testi 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

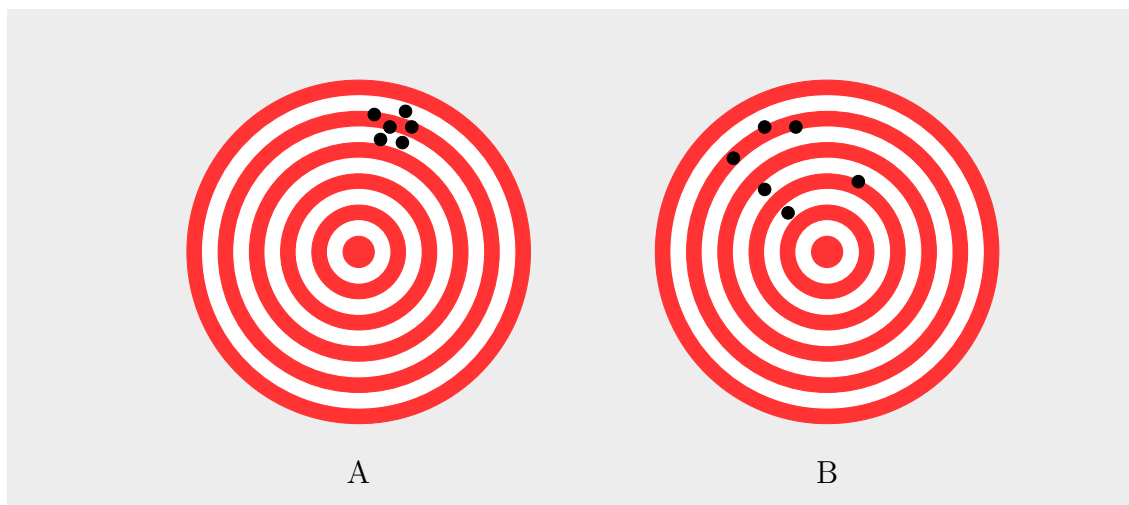
Arusaadavatel põhjustel on testi valiidsus äärmiselt oluline ning psühholoogiliste testide valiidsuse hindamine on olnud üks psühhomeetria põhilistest uurimisobjektidest. Sellega seoses on tehtud palju tööd ka valiidsuse definitsiooni täpsemaks muutmiseks, võrreldes selle paragrahvi alguses antud väga intuiitiivse 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:



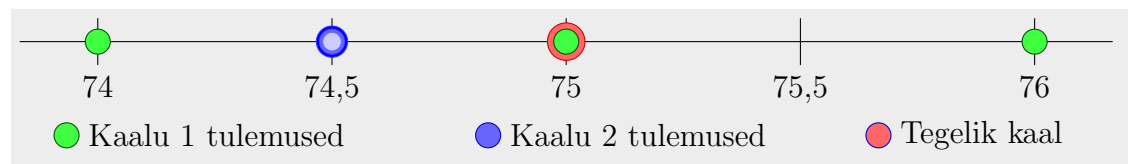
Joonis 4: Kahe küsimustiku reliaabluse võrdlus

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

Reliaabluse täpsema, matemaatilise definitsiooni anname edaspidi.

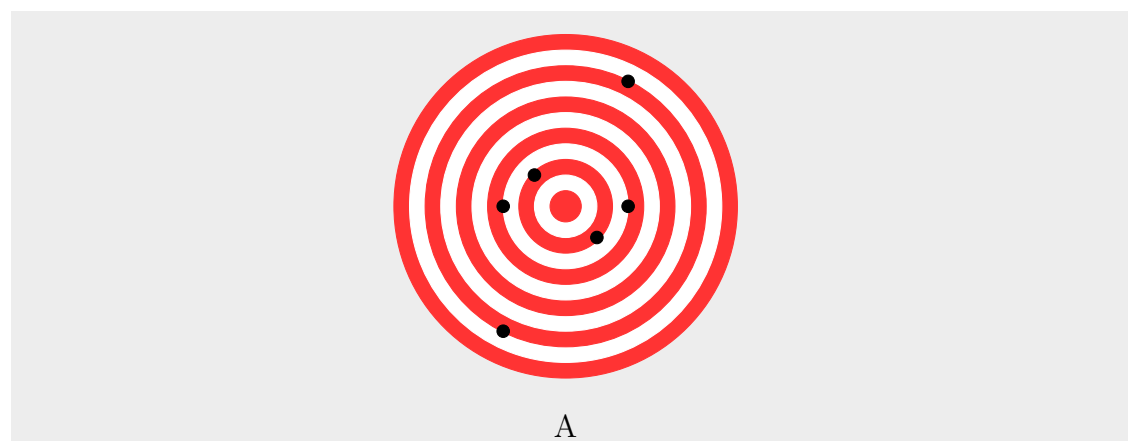
1.5 Valiidsuse ja reliaabluse suhe

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. Intuiitiivselt 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 testide – puhul on mingi va-
liidsuse taseme jaoks tarvilik tingimus mingi reliaabluse tase.

1.6 Sisemine järjepidevus

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

Sisemine järjepidavus (*internal consistency*) on küsimustiku karakteristik, mis ise-
loomustab testi erinevate küsimuste vastuste järjepidevust ehk seda, kui hästi on
vastused ühist konstruktsiooni hindavatele küsimustele kooskõlas [7, 177]. Pane-
me tähele, et erinevalt eelnevalt toodud reliaabluse definitsioonist, on sisemine
järjepidevus ühe testi 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 ole-
ma 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 [1], 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 [12] [11] ja selle tõlgenduse Klaas Sijtsma poolt. [14, 109]

2.1 Reliaabluse definitsioon

MEELDETULETUSEKS

Definitsioon. *Elementaarsündmuseks* nimetatakse juhusliku katse võimalikku tulemust.^a

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 $x_1, x_2, \dots (x_n)$.

^a Siin ja edaspidi toetume tõenäosusteooria definitsioonide ja tulemuste sõnastamisel Kalev Pärna õpikule [?].

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 keskväärtuse ehk

$$t_i = \epsilon[X_i].^{②}$$

Kuna juhusliku suuruse X_i võimalike väärtuste arv on lõplik, on ka tema keskväärtus 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 keskväärtus 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 keskväärtuse 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 $X' = T' + E'$, siis kehtib

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

Teisisõnu, küsimustiku tegeliku tulemuse ja küsimustiku mõõtmisvea vahel ei ole korrelatsiooni. See omadus muudab mudeli rakendamise palju lihtsamaks, kuna me ei pea eeldama, et meie koostatud küsimustikul on selline omadus. Kahjuks ei mahu selle omaduse tõestus käesoleva töö raamidesse, rangest tõestusest huvitatutele soovitame tutvuda Donald Zimmermani mudeli ülesehitusega [19].

Paneme tähele, et kuna me ei seadnud testidele X ja X' mingeid kitsendusi, siis ka $\text{cov}(T, E) = 0$.

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. 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

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

Paneme tähele, et kahe paralleelse küsimustiku 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

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

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

$$\rho_{XX'} = \frac{D[T]}{D[X]} = \frac{D[X - E]}{D[X]} = \frac{D[X] - D[E]}{D[X]} = 1 - \frac{D[E]}{D[X]}. \tag{2.10}$$

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

Kuna küsimustiku tegelik tulemus on suurus, mida me praktikas ei tea, ning ka paralleelse küsimustiku läbiviimine ei ole tihti võimalik, on kasutusele võetud erinevad meetodid testi reliaabluse hindamiseks küsimustiku ühekordse läbiviimise käigus saadud informatsiooni põhjal. Järgnevalt vaatleme, kuidas seda teha.

2.2 Vajalikud taustteadmised reliaabluse hindamiseks

Järgnev käsitus tugineb Jacksoni ja Anguwamba 1977. aasta artiklile. [8] 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 kovariatsioo-

nimaatriksid^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 maatriksi Σ_X elemente kui x_{ij} , maatriksi Σ_T elemente kui t_{ij} ning maatriksi 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 kovariatsioonimaatriksi elementide summaga, ehk kui $X = \sum_{i=1}^k$ ja Σ_X on vektori (X_1, X_2, \dots, X_K) kovariatsioonimaatriks, siis

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

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

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 sinisega on tähistatud $D[X_1]$, punasega 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 rohelisega 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.

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 X_1, X_2, \dots, X_n , kus $n \in \mathbb{N}$ ja olgu Σ_X selle vektori kovariatsioonimaatriks. Olgu meil suvaline k -vektor $v = (v_1, v_2, \dots, v_n)$. Konstrueerime uue juhuslike suuruste vektori $(X'_1, X'_2, \dots, X'_n)$, kusjuures $X'_i = v_i X_i$ ning olgu Σ'_X selle vektori kovariatsioonimaatriks. Vaatleme juhuslikku suurust $X' = \sum_{i=1}^n X'_i$. Kuna eelneva põhjal $D[X'] = \sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Xij}$ ja definitsiooni põhjal

$D[X'] \geq 0$, siis $\sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Xij} \geq 0$. Kuna Σ'_X on kovariatsioonimaatriks, siis

$$\sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Xij} = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(v_i X_i, v_j X_j) \stackrel{\text{Lause 2.15}}{=} \sum_{i=1}^n \sum_{j=1}^n v_i v_j \text{cov}(X_i, X_j). \quad (2.16)$$

Paneme tähele, et

$$\begin{aligned} v^T \Sigma_X v &= \\ &= \begin{pmatrix} v_1 \text{cov}(X_1, X_1) + v_2 \text{cov}(X_1, X_2) + \dots v_n \text{cov}(X_1, X_n) \\ v_1 \text{cov}(X_2, X_1) + v_2 \text{cov}(X_2, X_2) + \dots v_n \text{cov}(X_2, X_n) \\ \vdots \\ v_1 \text{cov}(X_n, X_1) + v_2 \text{cov}(X_n, X_2) + \dots v_n \text{cov}(X_n, X_n) \end{pmatrix} v = \\ &= \sum_{i=1}^n \sum_{j=1}^n v_i v_j \text{cov}(X_i, X_j) \stackrel{(2.16)}{=} \sum_{i=1}^n \sum_{j=1}^n \Sigma'_{Xij} = D[X'] \end{aligned}$$

Seega $v^T \Sigma_X v \geq 0$, ning kuna v oli valitud suvaliselt, siis Σ_X on mittenegatiivselt määratud. Kuna n ja X_1, X_2, \dots, X_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 [6]. Viienda vaadeldava tõkke pakkusid välja Jackson ja Anguwamba oma 1977. aasta artiklis [8]. 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 . Numberdame 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. Analooogiliselt 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 Σ_T . 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

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

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

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

Seega ei ole maatriks enam positiivselt määratud, mis on eelmises peatükis läbiviidud 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 $\sum_{i=1}^k \theta_i \leq \sum_{i=1}^k x_{ii}$. Võttes arvesse võrdust (2.14), oleme saanud küsimustiku, mille vaadeldav tulemus on X , reliaabluse alumise tõkke, mida tähistame kui λ_1 ,

kus

$$\lambda_1 = 1 - \frac{\sum_{i=1}^k x_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \leq 1 - \frac{\sum_{i=1}^k \theta_i}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{D[E]}{D[X]}. \quad (3.1)$$

3.2 Alumine tõke λ_2

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

Olgu meil mitte-negatiivselt määratud sümmeetriline maatriks $C, C \in Mat_n(\mathbb{R})$. Näitame, et sellisel juhul kehtib väide, et iga i, j korral

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

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 me soovisimegi veenduda. Eelneva arutelu idee pärineb Ivar Tammeraidi õpikust. [17].

Olgu meil mingid arvud $y_i, 1 \leq i \leq n$. Paneme tähele, et siis kehtib järgnev võrdus

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

Järelikult kehtib ka

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

Liidetavad teistmoodi grupeerides saame, et

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

Vaatleme jällegi kovariatsiooni maatriksit Σ_E . Märkame, et kuna $x_{ii} = \theta_i + t_{ii}$ iga $1 \leq i \leq k$ korral, siis kehtib järgnev:

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

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

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

Eelnevalt käsitletud algebralise samasuse ning kovariatsioonimaatriksite mitte-negatiivselt määratuse ja sümmeetrilisuse põhjal kehtib

$$\begin{aligned}
\sum_{i=1}^k t_{ii} &= \sqrt{\frac{\sum_{i=1}^k \left(\sum_{j=1}^{i-1} (t_{ii} - t_{jj})^2 + \sum_{m=i+1}^k (t_{ii} - t_{mm})^2 \right) + 2k \sum_{i=1}^k \left(\sum_{j=1}^{i-1} t_{ii} t_{jj} + \sum_{m=i+1}^k t_{ii} t_{mm} \right)}{2(k-1)}} \geq \\
&\geq \sqrt{\frac{k \sum_{i=1}^k \left(\sum_{j=1}^{i-1} t_{ii} t_{jj} + \sum_{m=i+1}^k t_{ii} t_{mm} \right)}{(k-1)}}
\end{aligned}$$

Arvestades, et kovariatsioonimaatriks on mitte-negatiivselt määratud ning seda, et kui $i \neq j$, siis $t_{ij} = x_{ij}$, saame eelnevalt tõestatud abitulemuse põhjal väita, et

$$\sum_{i=1}^k t_{ii} \geq \sqrt{\frac{k \sum_{i=1}^k \left(\sum_{j=1}^{i-1} t_{ii} t_{jj} + \sum_{m=i+1}^k t_{ii} t_{mm} \right)}{(k-1)}} \geq \sqrt{\frac{k \sum_{i=1}^k \left(\sum_{j=1}^{i-1} x_{ij}^2 + \sum_{m=i+1}^k x_{im}^2 \right)}{(k-1)}}$$

Selle taustal saame defineerida reliaabluse alumise tõe, mida tähistame λ_2 :

$$\lambda_2 = 1 - \frac{\sum_{i=1}^k x_{ii} - \sqrt{\frac{k \sum_{i=1}^k \left(\sum_{j=1}^{i-1} x_{ij}^2 + \sum_{m=i+1}^k x_{im}^2 \right)}{(k-1)}}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \leq 1 - \frac{\sum_{i=1}^k x_{ii} - \sum_{i=1}^k t_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{\sum_{i=1}^k \theta_i}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{D[E]}{D[X]}.$$

Paneme tähele, et

$$\lambda_2 = \lambda_1 + \frac{\sqrt{\frac{k \sum_{i=1}^k \left(\sum_{j=1}^{i-1} x_{ij}^2 + \sum_{m=i+1}^k x_{im}^2 \right)}{(k-1)}}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}}$$

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

$$2(k-1) \sum_{i=1}^k t_{ii} \geq 2 \sum_{i=1}^k \left(\sum_{j=1}^{i-1} x_{ij} + \sum_{m=i+1}^k x_{im} \right),$$

mis on samaväärne sellega, et

$$\sum_{i=1}^k t_{ii} \geq \frac{1}{k-1} \sum_{i=1}^k \left(\sum_{j=1}^{i-1} x_{ij} + \sum_{m=i+1}^k x_{im} \right).$$

Eelneva põhjal oleme saanud järgmise reliaabluse alumise tõkke:

$$\begin{aligned} \lambda_3 &= 1 - \frac{\sum_{i=1}^k x_{ii} - \frac{1}{k-1} \sum_{i=1}^k \left(\sum_{j=1}^{i-1} x_{ij} + \sum_{m=i+1}^k x_{im} \right)}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} \leq \\ &\leq 1 - \frac{\sum_{i=1}^k x_{ii} - \sum_{i=1}^k t_{ii}}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{\sum_{i=1}^k \theta_i}{\sum_{i=1}^k \sum_{j=1}^k x_{ij}} = 1 - \frac{D[E]}{D[X]}. \end{aligned}$$

Tõkke teisendamisel saame, et

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

Tuletades meelde tõkke λ_1 definitsiooni (3.1) on ilmne, et

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

ja seega $\lambda_1 \leq \lambda_3$. Lisaks on võimalik näidata, et $\lambda_3 \leq \lambda_2$, näiteks [13, 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 [2], ning see on ilmselt enim kasutatud reliaabluse hinnang. Käesolevas töös jääme selle alumise tõkke nimetamisel siiski λ_3 juurde – nii nimetas tuletatud tõket Louis Guttman oma kuus aastat varem ilmunud artiklis [6].

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

3.4 Alumine tõke λ_4

Vaatleme kovariatsioonimaatriksit Σ_T . 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 Σ_E 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 alumise tõkke λ_4 kui maksimumi kõikidest sellistest tõketest, seega

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

3.5 Alumine tõke glb

Paneme tähele, et Σ_E struktuuri (2.12) tõttu on tema mitte-negatiivselt määratus samaväärne sellega, et $0 \leq \theta_i$ iga $1 \leq i \leq k$ korral.

Tõepoolest, olgu Σ_E mitte-negatiivselt määratud, ning oletame vastuväiteliselt, et leidub $i, 1 \leq i \leq k$ nii, et $0 > \theta_i$. Olgu v k -vektor, nii et

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

Siis $v^T \Sigma_E v = \theta_i < 0$, mis on aga vastuolu sellega, et Σ_E on mitte-negatiivselt määratud. Teistpidi, kehtigu $0 \leq \theta_i$ iga $1 \leq i \leq k$ korral. Olgu v suvaline k -vektor. Siis $v^T \Sigma_E v = \sum_{i=1}^k v_i^2 \theta_i \geq 0$. Kuna v oli suvaline, siis on Σ_E vastavalt definitsioonile mitte-negatiivselt määratud.

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:

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

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

Olgu z eelneva optimeerimisülesande lahend Σ_X korral. Tähistame tähekombinatsiooniga *glb*, mis tuleb fraasist *greatest lower bound*, lahendi z abil saadud reliaabluse alumise tõkke, 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 . Samuti on raske leida empiirilist näidet, kus λ_4 ei ole tingimata suurem oleks väiksem 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 kovariatsoonimaatriksi korral on *glb* hinnangust λ_3 oluliselt suurem, vastavalt 0,77 ja 0,43. Joonisel 16 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. [13] [15] [3]

Mainime siinkohal ära, et hinnangud *glb* ja λ_4 võivad teatud tingimustes olla nihkega. See on probleem, mis esineb praktikas, kui kasutame kovariatsiooni maatriksi asemel tema valimi põhjal saadud hinnangut. Tõkke *glb* korral vaata näiteks [?], λ_4 puhul [?].

4 Programmiteek

Järgnevalt põhjendan lühidalt Pythoni eeliseid teegi realiseerimisel võrreldes selliste alternatiividega nagu Java või MATLAB:

1. Python on keel, millega programmeerimisalased õpingud tavaliselt algavad. Näiteks Tartu Ülikoolis õpetatakse programmeerimist keeles Python nii matemaatika-informaatika teaduskonna tudengitele suunatud aines kui ka teistele huvilistele suunatud aines "Programmeerimise 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.

Meetodid, mis leiavad λ_1 , λ_2 , λ_3 ja λ_4 , on implementeeritud nii programmeerimiskeeles Python 2 kui ka programmeerimiskeeles Python 3. Klass, mille abil saab leida *glb*, on implementeeritud sõltuvate teekide piirangute tõttu ainult programmeerimiskeeles Python 2. Programmiteegi loomisel on kasutatud teeki SciPy [10]

ja CVXPY [4]. Programmiteek on kättesaadav aadressil <https://github.com/vootelerotov/CTT-reliability-estimates>.

Lisas C on ära toodud hinnangute implementatsioonid programmeerimiskeeles Python 2.

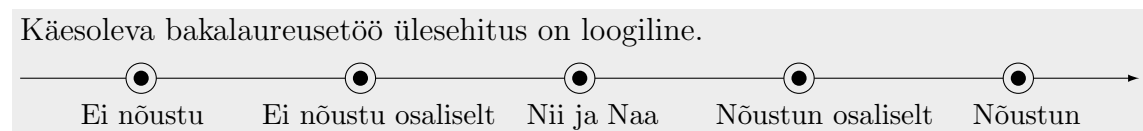
5 Intervallskaal alternatiivne tõlgendamine

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

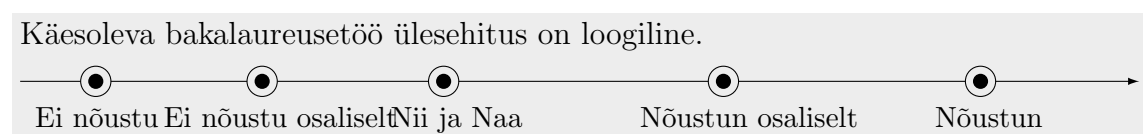
5.1 Vastuste tõlgendamine intervallskaalal

Käesoleva töö sissejuhatuses kirjeldasime viisi, kuidas vaikumisi 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 vaikumisi meetod? Vaatleme kahte näidet:



Joonis 9: Näide sellest, kuidas paigutuvad hinnangud skaalal vaikumisi meetodit kasutades



Joonis 10: Näide hinnangute alternatiivsest paiknemisest skaalal

Vaikumisi tõlgendusele alternatiive välja pakkudes peab meil olema mingi karakteristik, mille alusel tõlgendust hinnata. Esitame järgmise idee: üritame leida hin-

nangute paiknemist skaalal nii, et küsimustiku reliaablus oleks võimalikult suur. Seame endale loomuliku piirangu: vastuste esialgne järjestus ei tohi muutuda.

5.2 Reliaabluse hindamine

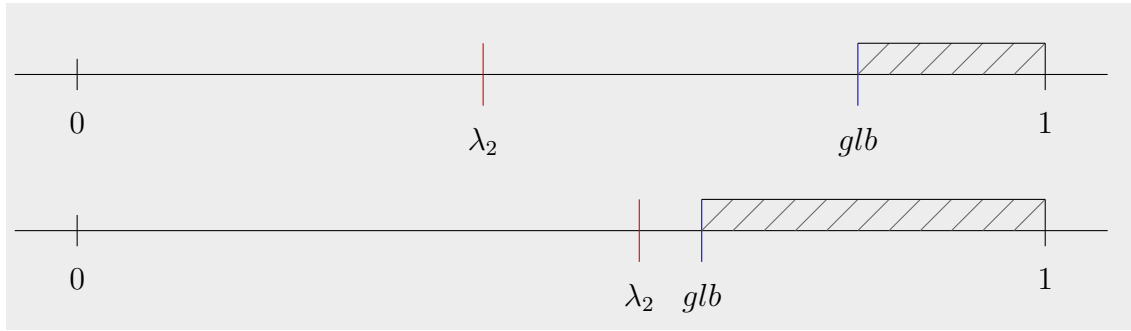
Kuna küsimustiku reliaabluse leidmine ei ole praktiliselt võimalik, siis kasutame selle lähendusena eelmises peatükis käsitletud alumisi tõkkeid. On selge, et parima tulemuse annab võimalikult täpse alumise tõkke kasutamine. Kuna sobiva tõlgenduse leidmine taandub optimeerimisülesandele, siis kasutame 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 11: 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 tõe on vaikumisi tõlgenduse korral suurem kui leitud tõlgenduse korral, siis jääda vaikumisi 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 valiidsust? Selle küsimuse korral märgime ära, et kuigi vaikumisi 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 valiidsus parem kui teise korral. Ühelt poolt on selline vastus mitterahuldav – meie pakutud lahendus ei pruugi olla parem kui olemasolev. Teisalt ei pruugi vaikumisi 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 tõe 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 naturaalarvu-

dega 1 kuna 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 vaikumisi 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 $M_1, M_2, M_2, \dots, M_n$, kusjuures juhusliku suuruse M_i määramispiirkond ühtib suuruse K_i määramispiirkonnaga 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 kitsendused:

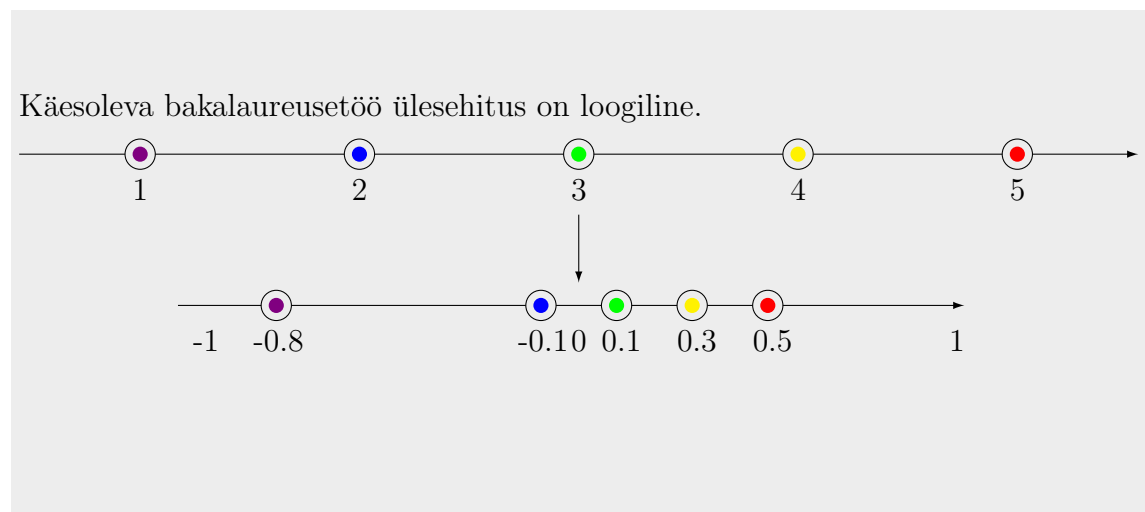
$$E(L_i) = \sum_{j=1}^l p_{ij} \mu_{ij} = 0, \quad (5.1)$$

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

Paneme tähele, et selliste piirangute korral esitub λ_3 lihtsamal kujul. Tõepoolest,

$$\lambda_3 = \frac{k}{k-1} \left(1 - \frac{\sum_{i=0}^k D(M_i)}{\sum_{i=1}^k \sum_{j=1}^k \text{cov}(M_i, M_j)} \right) = \frac{k}{k-1} \left(1 - \frac{k}{\sum_{i=1}^k \sum_{j=1}^k \text{cov}(M_i, M_j)} \right).$$

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 12: 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

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

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

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

Eelneva põhjal piisab meile, arvestades eelnevalt äratoodud piiranguid L_i keskväärtusele ja dispersioonile, 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 optimiseerimisülesande lahendamine. Kahjuks on ülesanne sellist tüüpi, mille lahendamine on üldjuhul raske^④. Küll on sellist tüüpi ülesandeid võimalik lõdvendada, levinud lõdvenduseks on probleemi teisendamine mittenegatiivseks optimiseerimisülesandeks [?, 2] – see lõdvendus osutus antud probleemi puhul aga liiga lõdvaks.

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

^④ Täpsemalt on tegemist *NP-Hard* probleemiga: kui sellele ülesandele leidub polünoomiaalse keerukusega lahendusalgorithm, siis leidub polünoomiaalse keerukusega algoritm ka paljudele teistele 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] Lee J. Cronbach. Test "Reliability": Its Meaning and Determination. *Psychometrika*, 12(1):1–16, 1947.
- [2] Lee J. Cronbach. Coefficient Alpha and the Internal Structure of Tests. 16(3):297–330, 1951.
- [3] Lee J. Cronbach and Richard J. Shavelson. My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement*, 64(3):391–418, June 2004.
- [4] S. Diamond, S; Chu, E; Boyd. CVXPY: A Python-Embedded Modeling Language for Convex Optimization, 2014.
- [5] Diane R. Edmondson. Likert Scales: A History. In *Conference on Historical Analysis & Research in Marketing Proceedings*, 2005.
- [6] Louis Guttman. A Basis for Analyzing Test-Retest Reliability. *Psychometrika*, 10(4):255–282, 1945.
- [7] Robin K. Henson. Understanding Internal Consistency Reliability Estimates: A Conceptual Primer on Coefficient Alpha. *Measurement & Evaluation in Counseling & Development*, 34(3):177, 2001.
- [8] Paul H Jackson and Christian C. Agunwamba. Lower Bounds for the Reliability of the Total Score on a Test Composed of Non-Homogeneous Items: I: Algebraic Lower Bounds. *Psychometrika*, 42(2):567–578, 1977.
- [9] Susan Jamieson. Likert scales: how to (ab)use them. *Medical education*, 38(12):1217–8, December 2004.
- [10] Eric Jones, Travis Oliphant, and Pearu Peterson. SciPy: Open Source Scientific Tools for Python.
- [11] F.M. Lord and M.R Novick. *Statistical theories of mental test scores*. 1968.
- [12] Melvin R. Novick. The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1):1–18, February 1966.
- [13] William Revelle and Richard E. Zinbarg. Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1):145–154, December 2008.
- [14] Klaas Sijtsma. On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, 74(1):107–120, 2009.

- [15] Klaas Sijtsma. Reliability Beyond Theory and Into Practice. *Psychometrika*, 74(1):169–173, 2009.
- [16] Nicholas Smith. Example of Likert Scale. http://en.wikipedia.org/wiki/File:Example_Likert_Scale.svg.
- [17] Ivar Tammeraid. *Lineaaralgebra rakendused*. TTÜ Kirjastus, Tallinn, 1999.
- [18] UT. Õisi tagasiside. <http://is.ut.ee>.
- [19] D. W. Zimmerman. Probability spaces, hilbert spaces, and the axioms of test theory. 40(3):395–412, 1975.

Lisa

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

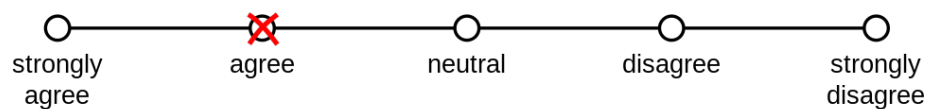
5.1. Palun kommenteerige ja/või põhjendage oma hinnangut.

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

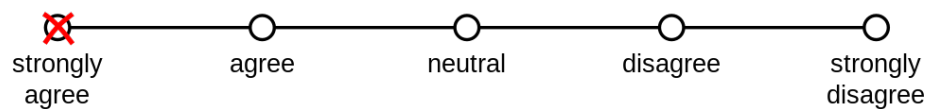
☐ nõustun igati ☐ pigem nõustun ☐ pigem ei nõustu ☐ ei nõustu üldse ☐ nii ja naa

Joonis 13: Näide Tartu Ülikooli õppeinfo süsteemi tagasiside ankeedist, kus rakendatakse Likerti skaalat [18]

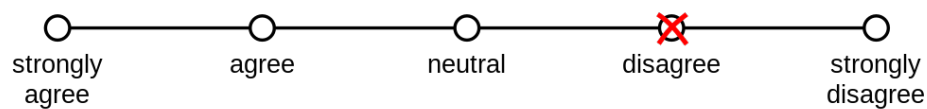
1. Wikipedia has a user friendly interface.



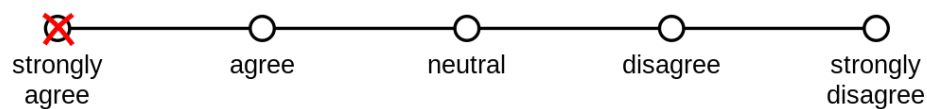
2. Wikipedia is usually my first resource for research.



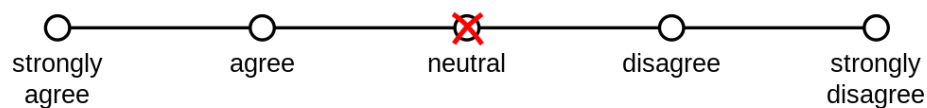
3. Wikipedia pages generally have good images.



4. Wikipedia allows users to upload pictures easily.



5. Wikipedia has a pleasing color scheme.



Joonis 14: Näide küsimustikust, kus on rakendatud Likerti skaalat; küsimused on paigutatud nende järjestikulisuse rõhutamiseks teljele [16]

B Empiirilised näited

Järgnevad kovariatsioonimaatriksid on leitud ühe psühholoogilise 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 15: Valimi 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 | glb |
|---------|-------------|-------------|-------------|-------------|-------|
| Väärtus | 0.79 | 0.87 | 0.86 | 0.93 | 0.93 |

Joonis 16: Valimi 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 | glb |
|---------|-------------|-------------|-------------|-------------|-------|
| Väärtus | 0.39 | 0.53 | 0.43 | 0.68 | 0.70 |

Joonis 17: Valimi 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     return 1 - (sum_of_main_diagonal()/
13                 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    n = len(cov_matrix)
12    l1 = calculate_lambda_1(cov_matrix)
13    sum_of_squares = sum_of_squares_of_non_diagonal_matrix_elements()
14    under_square_root = ((n)/(n-1)* sum_of_squares)**(1/2)
15    return l1 + (under_square_root/
16        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     def objective_function(u):
13         return u.T.dot(cov_matrix.dot(u))

14     def try_vectors():
15         u"""
16         Idea : generate half of all possible vectors of length n,
17         such that if vector v is in, vector -v is not.
18         Using binary representation as string,
19         from 2**(n-1) to 2**n, coding "0" to "-1"
20         """

```

```

21     smallest = np.Infinity
22     result_vector = []
23     l = []
24     for i in xrange(2**(n-1), 2**n):
25         binary_of_i = np.binary_repr(i, width=n)
26         binary_of_i_int = [ 1 if x == u"1" else -1 for x in binary_of_i]
27         u = np.array(binary_of_i_int)
28         result = objective_function(u)
29         if result < smallest:
30             smallest = result
31             result_vector = u
32     return smallest

33 def calc_lambda(smallest):
34     return 1 - (smallest/
35         util.calculate_sum_of_covariance_matrix_elements(cov_matrix, n))

36 n = util.check_that_matrix_is_square_and_fix_n(cov_matrix)
37 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, 2. juuni 2014. a.