Bakatöö

Vootele Rõtov

9. mai 2014. a.

Sisukord

1 Sissejuhatus

Käesoleva bakalaureuse töö on saanud motivatsiooni psühholoogide praktilisest probleemist - kuidas hinnata nende töövaldkonnas tihti kasutatavate valikvastustega testide headust. Tuleb välja, et ühe testide headust kirjeldava karakteristiku - reliaabluse, hindamiseks sobivad vahendid saame klassikalise testiteooria (*Classical Test Theory*) nimelise matemaatilise teooria tulemuste põhjal.
Töö esimeses osas anname vajaliku taustinfo ja probleemipüstituse. Seejärel ehitame üles meile vajamineva osa klassikalisest testiteooriast - defineerime reliaabluse ning rajame vundamendi selle hindamiseks. Järgneval vaatleme erinevaid võimalusie reliaabluse hindamiseks. Lõpetuseks uurime põgusalt ühte võimalikku lähenemist järjestikskaala tõlgendamiseks intervallskaalana, mis kasutab ära eelnevalt uuritud reliaabluse hinnanguid ja sättestab tõlgendamis probleemi optimiseerimisprobleemina.
Töö eesmärgiks on anda eesti keelne ülevaade klassikalisest testiteooria alustest, mida autorile teadaolevalt ei ole käsitletud matemaatilisest aspektist vaadelduna - olemasolevad käsitlused on mõeldud praktiliseks abivahendiks psühholoogile ning seetõttu teistsuguse lähenemisega. Peale selle on töö eesmärgiks ka reliaabluse hindamiseks sobivate vahendite tutvustamine tasemel, millest võiks loodetavasti kasu olla ka teste koostavatele praktikutele. Uurimuse viimases osas loodame lugejale tutvustada tavapärasest teistsuguse lähenemist testi tulemuste kvantitiivsele tõlgendamisele, mis tugineb viimase aja edusammudele optimiseerimise valdkonnas.
Lisaks loetavale dokumendile on töö väljundiks ka teek, mis sisaldab funktsioone kõigi töös välja toodud hinnangute leidmiseks testi tulemuste põhjal ning leiab loodetavasti psühholoogide poolt kasutamist.

2 Taustinfo ja probleemipüstitus

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

2.1 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 vastuse variantide 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

2.2 Likerti skaala tõlgendamine intervallskaalana

Likerti skaala tõlgendamisel intervallksaalana on väljakujunenud tava seada valikvastustele vastavusse järjestatud täisarvud, kusjuures mida positiivsem vastuseevariant, seda suurem 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 invtervallskaalana

Lugejal võib tekkida õigustatud küsimus, kuidas põhjendab autor Likerti skaala käsitlemist intervallskaalana, kui Likerti skaala on olemuseslt järjestikskaala ning

[1]Näited terviklikest küsimustikest on lisades, joonisel 9 ja 10

selle tõlgendamises intervallskaalana on vastuoluline küsimus, näiteks [10]. Siinkohal tõdeme, et Likerti skaala tõlgendamine intervallksaalana on praktikas piisavalt levinud, et selle valdkonna uurimine õigustatud oleks - olenamat selle teoreetilisest põhjendatusest. Siinkohal väärib autori silmis esile toomist kriitikute üks levinumaid argumente - „hea“ ja „väga hea“ keskmine ei ole loomulikul viisil tõlgendatav kui „hea + pool“ ehk pole mingit põhjust eeldada, et kõikide küsimuste omavaheline kaugus on mingil põhjusel võrdne. Autor nõustub selle kriitikaga ning loodab töö neljandas osas pakkuda alternatiivse tõlgendusviisi.

2.3 Testi valiidsus

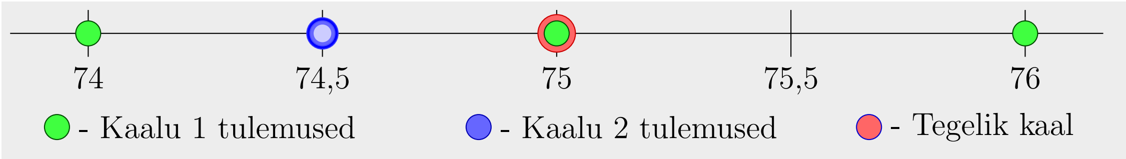
Testi **valiidsus** on testi karakteristik, mis iseloomustab testi võimet mõõta seda, mida ta disainiti mõõtma. Enamikes käsitlustes vaadeltakse valiidsust kui väärtust intervall skaalal nulli ja ühe vahel.
Näitena, kaal mis näitab 75 kilo kaaluva inimese kaaluks 74,5 kilo omab kõrgemat valiidsust kui kaal, mis sama inimese puhul näitab kaaluks 65 kilo.
Arusaadavatel põhjustel on testi valiidsus äärmiselt oluline ning psühholoogiliste testide valiidsuse hindamine on olnud üks psühhomeetria põhilistest uurimisobjektidest. Sellega seoses on palju tööd tehtud ka valiidsuse definitsiooni täpsemaks muutmisega, võrreldes selle paragrahvi alguses antud väga intuitiivse definitsiooniga. Kuna valiidsus on selle töö raames vajalik vaid taustinfona probleemi mõttestamisel tundub antud lihtne definitsioon aga sobivaim.

2.4 Testi reliaablus

Testi **reliaablus** on testi omadust saada sama subjekti erinevatel mõõtmistel sama testiga sama tulemus ehk testi stabiilsus. Nagu valiidsus, on ka reliaablus reeglina määratud intervall skaalal, nulli ja ühe vahel.
Näitena, kaal mis näitab 75 kilo kaaluva inimese kaaluks 74,5 kilo igal kaalumisel omab kõrgemat reliaablust kui kaal, mis näitab juhuslikult kas 74,5 kilo või 75,5 kilo.
Reliaabluse täpsema, matemaatilise definitsioone anname edaspidi.

2.5 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 jättavad selle küsimuse lahtiseks - vaatleme järgmist olukorda:



Joonis 3: Kaalumiste tulemused kahe erineva kaaluga kolmel katsel

Kui mõista valiidsust kui testi keskmist tulemust, siis võime väita, et esimene kaal on väiksema reliaabluse aga suurema valiidsusega. Selline valiidsuse mõttestamine aga ei ole küsimustike modeleleerimise korral otstarbekas - katsete kordamine on enamasti keeruline ning väheste mõõtmistulemuste põhjal ei ole võimalik tegeliku väikese reliaablusega testi tulemust välja selgitada.
Seega, meie vaadeldava olukorra - psühholoogiliste testide - korral on mingi valiidsuse taseme jaoks tarvilik tingimus mingi reliaabluse tase.

2.6 Sisemine järjepidevus

Reliaabluse kui termini probleemiks on tema mitmetähenduslikus. Toome siinkohal ära ühe mõiste, mida tihti samuti reliaablusena tuntakse.
Sisemine järjepidavus (*internal consistency* on testi karrakteristik, mis iseloomustab testi erinevate küsimuste vastuste järjepidevust ehk seda, kui hästi on kooskõlas ühist konstruktsiooni hindavad küsimused [8, 177]. Piltlikult väljendudes, olgu meil järgnev küsimustik:

Joonis 4: Küsimustik bakalaureusetöö ülesehituse kohta

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

Selle karrakteristiku mõõtmiseks 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 karakterisitku eristamine aitab lugejal käesolevas töös kergemini orienteeruda.
Märgime, et lisaks toodud kahele definitsioonile on ka teisi reliaabluse definitsioone [1], mis aga selle töö kontekstis ei tohiks segadust tekitada ning mille äratoomist siin ei ole autor pidanud vajalikus.

2.7 Probleemi püstitus

Käesolevo tööga üritame pakkuda lahendusti kahele probleemile. Esimene nendest seisneb testi reliaabluse hindamises - kas on võimalik testi tulemuste põhjal hinnata selle testi reliaablust? Teine probleem lähtub esimesest - kui reliaabluse hindamine on võimalik, kas siis on võimalik kasutada saadud hinnangut Likerti skaala paremaks tõlgendamiseks intervallskaalana? Esimese probleemi rangeks mõtestamiseks vaatleme järgnevalt reliaablust ühe matemaatilise raamisitku - klassikalie testiteooria - kontekstis.

3 Reliaablus klassikalises testiteoorias

Järgevas võtame aluseks Melvin Novicki klassikalise testiteooria [13] [12] ja selle tõlgenduse Klaas Sijtsma poolt. [14, 109]

3.1 Reliaabluse definitsioon

MEELDETULETUSEKS

Definitsioon. *Elementaarsündmuseks* nimetatakse juhusliku katse võimaliku tulemust.

Definitsioon. *Tõenäosuseks* nimetatakse funktsiooni P, mis igale sündmusele $A \in 2^{\Omega}$, kus Ω on mingi juhusliku katse kõikvõimalikke 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 $A_1, A_2, ...$ on teineteist välistavad, siis
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$$
.

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

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 kõike erinevaid võimalikke vastuste kombinatsioone küsimustikule kui elementaarsündmusi, tähistame saadud elementaarsündmuste hulga kui Ω ning elementaarsündmusi kui ω. Eeldame ka tõenäosuste $P_i : 2^{\Omega} \rightarrow \mathbb{R}$ olemasolu, mida tõlgendame kui testi tulemuste esinemissagededust vastaja i korral.

Küsimus / ω	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 5: Näited elementaarsündmustest

Olgu X_i^j hulgal Ω määratud juhuslik suurus, mille võimalike 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 tõkkestatud arvuga l ning liidetavaid on k, siis ei saa funktsiooni muutumispiirkonna võimsus olla suurem kui l^k.

MEELDETULETUSEKS

Definitsioon. Disktreetse juhusliku suuruse X *jaotuseks* nimetatakse paaride komplekti (x_i, p_i) , $i = 1, 2, ...$, 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 $k_i = (x_i, p_i)$, $i = 1, 2, ...$, *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 $DX = \epsilon\left[(X - \epsilon X)^2\right]$.

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

$$t_i = \epsilon\left[X_i\right].^{②}$$

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 E_i. Seega $X_i = t_i + E_i$.

MEELDETULETUSEKS

Lause 3.1. Kui X on diskreetne juhuslik suurus, jaotusega K, kus K on paaride komplekt $k_i = (x_i, p_i), i = 1, 2, ...$ ning c on mingi konstant siis $\epsilon(X + c) = \epsilon(X) + c$

Tõestus.

$$\epsilon(X + c) = \sum_{\{i|k_i \in K\}} (x_i + c)p_i = \sum_{\{i|k_i \in K\}} x_i p_i + \sum_{\{i|k_i \in K\}} cp_i =$$
$$= \sum_{\{i|k_i \in K\}} x_i p_i + c \sum_{\{i|k_i \in K\}} p_i = \sum_{\{i|k_i \in K\}} x_i p_i + c = \epsilon(X) + c$$

□

Lause 3.2. Juhuslikule suurusele konstandi liitmine ei muuda dispersiooni : $D(X + c) = DX$

Tõestus.

$$D(X+c) = \epsilon[X+c-\epsilon(X+c)] \overset{\text{Lause 3.1}}{=} \epsilon[X+c-\epsilon(X)+c] = \epsilon[X-\epsilon(X)] = DX$$

□

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

$$t_i = \epsilon\left[X_i\right] = \epsilon\left[t_i + E_i\right] \overset{\text{Lause 3.1}}{=} t_i + \epsilon\left[E_i\right] \implies \epsilon\left[E_i\right] = 0.$$

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

$$D\left[X_i\right] = D\left[E_i + t_i\right] \overset{\text{Lause 3.2}}{=} D\left[E_i\right].$$

②Tähistame keskväärtuse tähega ϵ kuna tavapärane tähistus E leiab antud käsitluses teistsuguse rolli.

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

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

kus funktsiooni P_p tõlgendame kui isiku $i \in I$ ja testi tulemuse $\omega \in \Omega$ koos esinemis sagedust.
Paneme tähele, et sellel elemntaarsündmuste hulgal võime määrata diskreetse juhusliku suuruse T, nii et

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

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

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

ja testi vea kui diskreetse juhusliku suuruse E, nii et

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

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

$$X\left((i,\omega)\right) = X_i(\omega) \overset{def}{=} t_i + E_i(\omega) = T\left((i,\omega)\right) + E\left((i,\omega)\right).$$

Eeldame edaspidises, et X ei ole konstante juhuslik suurus - kui X on konstante, siis .

MEELDETULETUSEKS

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

$$cov(X,Y) = \epsilon\left[(X - \epsilon X)(Y - \epsilon Y)\right].$$

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

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

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

$$cov(T, E') = 0, \tag{3.3}$$

ehk mingi küsimustiku tegeliku tulemuse ja mingi küsimustiku mõõtmisvea vahel ei ole korrelatsiooni. See omadus muudab mudeli rakendamise palju lihtsamaks, kuna me ei pea eeldame, et meie poolt koostatud küsimustikul on selline omadus. Tõestus sellele omadusele kahjuks käesoleva töö raamidesse ei mahu, rangest tõestusest huvitatutele soovitame tutvuda Donald Zimmermani mudeli ülesehitusega [**?**].
Paneme tähele, et kuna me ei seadunud mingeid kitsendusi testide X ja X' jaoks, siis ka $cov(T, E) = 0$.
Küll aga peame me tegema järgneva eelduse - olgu meil kaks erinevat küsimustiku, mille mõõtmisvead on vastavalt E ja E' siis

$$cov(E, E') = 0. \tag{3.4}$$

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

MEELDETULETUSEKS

Lause 3.5. Juhuslikke suuruste X ja Y puhul kehtib

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

Tõestus.

$$D(X+Y) \overset{def}{=} \epsilon[X+Y-\epsilon(X+Y)]^2 \overset{\epsilon \text{ lin.}}{=} \epsilon[X - \epsilon X + Y + \epsilon Y]^2 =$$
$$= \epsilon(X-\epsilon X)^2 + 2\epsilon(X-\epsilon X)(Y-\epsilon Y) + \epsilon(Y-\epsilon Y)^2 \overset{\text{cov def}}{=}$$
$$= D(X) + D(Y) + 2cov(X, Y)$$

□

Järgnevalt tutvume paralleelse küsimustiku mõistega. Kaks küsimustiku, 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

$$T = T'. \tag{3.6}$$

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 paraleelse küsimustiku X ja X' korral kehtib

$$D[X] = D[X'], \tag{3.7}$$

kuna

$$D[X] = D[T + E] \stackrel{\text{Lause 3.5}}{=} D[T] + D[E] + 2cov(T, E) \stackrel{(3.3)}{=} D[T] + D[E] =$$
$$= D[T'] + D[E'] = D[T'] + D[E'] + 2cov(T', E') = D[T' + E'] = D[X'].$$

MEELDETULETUSEKS

Lause 3.8. Olgu meil juhuslikud suurused X, Y ja Z. Siis

$$cov(X + Y, Z) = cov(X, Z) + cov(Y, Z).$$

Tõestus.

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

□

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

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

Eelneva põhjal saame näidata, et kehtib

$$\begin{aligned}
\rho_{XX'} &\overset{def}{=} \frac{cov(X,X')}{\sqrt{D\left[X\right]D\left[X'\right]}} \overset{(3.7)}{=} \frac{cov(X,X')}{\sqrt{D\left[X\right]D\left[X\right]}} = \frac{cov(T+E,T'+E')}{D\left[X\right]} \overset{\text{Lause 3.8}}{=} \\
&= \frac{cov(T,T')+cov(T,E')+cov(E,T')+cov(E,E')}{D\left[X\right]} \overset{(3.3)}{=} \\
&= \frac{cov(T,T')+cov(E,E')}{D\left[X\right]} \overset{(3.4)}{=} \frac{cov(T,T')}{D\left[X\right]} \overset{(3.6)}{=} \frac{cov(T,T)}{D\left[X\right]} \overset{def}{=} \\
&= \frac{\epsilon\left[(X-\epsilon\left[X\right])(X-\epsilon\left[X\right])\right]}{D\left[X\right]} \overset{def}{=} \frac{D\left[T\right]}{D\left[X\right]}.
\end{aligned} \tag{3.9}$$

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

$$\rho_{XX'} = \frac{D\left[T\right]}{D\left[X\right]} = \frac{D\left[X-E\right]}{D\left[X\right]} = \frac{D\left[X\right]-D\left[E\right]}{D\left[X\right]} = 1 - \frac{D\left[E\right]}{D\left[X\right]}. \tag{3.10}$$

Paneme tähele, kuna $D[X] = D[T] + D[E]$, siis eelneva põhjal $\rho_{XX'} \in [0,1]$.
Eelneva põhjal on selge, et väide “testi reliaablus on suurem” on samaväärne järgmise kolme väitega:

1. testi korrellatsioon paralleelse testiga on suurem,
2. testi tegeliku tulemuse dispersioon on suurem võrreldes testi vaadeldava tulemuse dispersiooniga,
3. testi mõõtmisvea dispersioon on väiksem võrreldes testi vaadeldava tulemuse dispersiooniga.

Märgime, et klassikalise testiteooria raames defineeritud reliaablus sobib kokku eelmises peatükkis antud mitteformaalse definitsiooniga, kuna küsimustik on iseendaga paralleelne testiks võib lugeda ka sama testi uuesti läbiviimise sama vastajate kogumi peal. Paneme tähele, et tegemist on suhteliselt nõrku eeldusi vajava teooriaga, kõigi raskem on praktikas süstemaatiliste vigade puudumist.
Kuna testi tegelik tulemus on suurus, mida me prakikas ei tea ning ka paralleelsete testide läbiviimine ei ole tihti võimalik on kasutusele on üritatud on võimalus testi reliaablust hinnata meie jaoks oluline. Järgnevalt vaatleme, kuidas seda teha.

3.2 Vajalikud taustteadimesd reliaabluse hindamiseks

Järgnev käsitlus tugineb Jacksoni ja Anguwamba 1977. aasta artiklile. [9] Olgu meil vastajate populatsioon I, üle mille me vaatleme küsimustiku, milles on k küsimust. Vaatleme seda küsimustiku klassikalise testiteooria raames ning olgu X küsimustiku vaadeldav tulemus, T küsimustiku tegelik tulemus ning E

küsimustiku mõõtmisviga. Vaatleme iga küsimust kui ühest küsimusest koosnevat alamküsimustiku ja tähistama fikseeritud küsimuse $i \in \{1, 2, \ldots, k\}$ vaadeltavat tulemust kui X_i, tegeliku tulemust kui T_i ning mõõtmisviga 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 peatükkis tähistasime alaindeksiga küsimusele vastajat siis edaspidi tähistama sellega kindlat küsimust. Vaatleme testi küsimuste tulemuste ja mõõtmisvea vektoreid: $(X_1, X_2, X_3, \ldots, X_n)$, $(T_1, T_2, T_3, \ldots, T_n)$, $(E_1, E_2, E_3, \ldots, E_n)$. Tähistame nende vektorite kovariatsioonimaatriksid[3], vastavalt $\Sigma_X, \Sigma_T, \Sigma_E$. Seega,

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

Eelnevalt läbiviidud arutelu (vt. (3.9)) põhjal teame, et $cov(X_i, X_j) = cov(T_i, T_j) + cov(E_j, E_j)$. Seega,

$$\Sigma_X = \Sigma_T + \Sigma_E. \tag{3.11}$$

Lisaks kehtib eelduse (3.4) põhjal saame, et et juhul kui $i \neq j$, siis $cov\,(E_i, E_j) = 0$. Sellest tulenevalt kehtib

$$\Sigma_E = \begin{pmatrix} cov\,(E_1, E_1) & 0 & 0 & \cdots & 0 \\ 0 & cov\,(E_2, E_2) & 0 & \cdots & 0 \\ 0 & 0 & cov\,(E_3, E_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & cov\,(E_n, E_n) \end{pmatrix}. \tag{3.12}$$

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

Tuletame meelde, et eelmises peatükkis veendusime (3.10), et testi reliaablus võrdub avaldisega $1 - \frac{D[E]}{D[X]}$.

Veendume, et juhuslike suuruste summa summa dispersioon võrdub liidetavate kovariatsioonimaatriksi elementide summaga ehk kui $X = \sum_{i=1}^{k}$ ja Σ_X on vektori

[3]Praktikas ei ole kovariatsioonimaatriksid meil teada, kuid meil on võimalik valimi põhjal saada hinnang vaadeldavate tulemuste kovariatsioonimaatriksile

$(X_1, X_2, \ldots, X_K)$ kovariatsioonimaatriks siis

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

Eelnevas veendumiseks paneme tähele, et

$$D[X] = D[X_1 + X_2 + \ldots + X_k] \overset{\text{Lause 3.5}}{=}$$
$$= D[X_1] + D[X_2 + X_3 + \ldots + X_k] + 2cov(X_1, X_2 + X_3 \ldots + X_k) \overset{\text{Lause 3.8}}{=}$$
$$= D[X_1] + D[X_2 + X_3 + \ldots + X_k] +$$
$$+ 2[cov(X_1, X_2) + cov(X_1, X_3) + \ldots + cov(X_1, X_k)].$$

Näeme, et tegemist on rekkurentse võrrandiga. Arvestades, et $cov(X, X) = D[X]$ (vt.(3.9)), 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 $2[cov(X_1, X_2) + cov(X_1, X_3) + \ldots + cov(X_1, X_k)]$ ja rohelisega liikmed, mille summast moodustub $D[X_1 + X_2 + \ldots + X_k]$. Siinkohal meenutame, et kovariatsioon on sümmeetriline ning seega $cov(X_i, X_j) = cov(X_j, X_i)$.

$cov(X_1, X_1)$	$cov(X_1, X_2)$	$cov(X_1, X_3)$	$\ldots$	$cov(X_1, X_k)$
$cov(X_2, X_1)$	$cov(X_2, X_2)$	$cov(X_2, X_3)$	$\ldots$	$cov(X_1, X_k)$
$cov(X_3, X_1)$	$cov(X_3, X_2)$	$cov(X_3, X_3)$	$\ldots$	$cov(X_3, X_k)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$cov(X_k, X_1)$	$cov(X_k, X_2)$	$cov(X_k, X_3)$	$\ldots$	$cov(X_k, X_k)$

Lisaks on sellel rekurentses võrandil on lihtne baasjuht, selleks on olukord, kus summa koosneb ainult ühest liidetavast.
Arvestades maatriksi 3.12 ehitust, saame võrduse

$$1 - \frac{D\,[E]}{D\,[X]} = 1 - \frac{\sum_i \theta_i}{\sum_i \sum_j x_{ij}}. \tag{3.13}$$

MEELDETULETUSEKS

Definitsioon. Maatriks $C \in Mat_k(\mathbb{R})$ on *mittenegatiivselt määratud* kui iga maatriksi $v \in Mat_{k,1}(\mathbb{R})$ 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 3.14. Kui X ja Y on juhuslikud suurused ja $\alpha \in \mathbb{R}$, siis

$$cov(X, \alpha Y) = \alpha cov(X, Y)$$

Tõestus.

$$cov(X, \alpha Y) = \epsilon\left[\left(X - \epsilon\left[X\right]\right)\left(\alpha Y - \epsilon\left[\alpha Y\right]\right)\right] \overset{\epsilon \text{ lin.}}{=}$$
$$= \epsilon\left[\left(X - \epsilon\left[X\right]\right)\left(\alpha Y - \alpha\epsilon\left[Y\right]\right)\right] = \epsilon\left[\alpha\left(X - \epsilon\left[X\right]\right)\left(Y - \epsilon\left[Y\right]\right)\right] =$$
$$= \alpha\epsilon\left[\left(X - \epsilon\left[X\right]\right)\left(Y - \epsilon\left[Y\right]\right)\right] = \alpha cov(X, Y)$$

□

Olgu meil suvaliste juhuslike suuruste vektor $X_1, X_2, \ldots, X_n$, kun $n \in \mathbb{N}$ ja olgu Σ_X selle vektori kovariatsioonimaatriks. Olgu meil suvaline k-vektor $v = (v_1, v_2, \ldots, v_n)$. Konstrueerime uue juhuslike suuruste vektori $(X'_1, X'_2, \ldots, X'_n)$, kusjuures $X'_i = v_i X_i$ ning olgu Σ'_X selle vektori kovariatsioonimaatriks. Vaatleme juhusliku 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,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \Sigma'_{Xij} = \sum_{i=1}^{n}\sum_{j=1}^{n} cov(v_i X_i, v_j X_j) \overset{\text{Lause 3.14}}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} v_i v_j cov(X_i, X_j). \quad (3.15)$$

Paneme tähele, et

$$v^T \Sigma_X v =$$
$$= \begin{pmatrix} v_1 cov(X_1, X_1) + v_2 cov(X_1, X_2) + \ldots v_n cov(X_1, X_n) \\ v_1 cov(X_2, X_1) + v_2 cov(X_2, X_2) + \ldots v_n cov(X_2, X_n) \\ \vdots \\ v_1 cov(X_k, X_1) + v_2 cov(X_k, X_2) + \ldots v_n cov(X_n, X_n) \end{pmatrix} v =$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} v_i v_j cov(X_i, X_j) \overset{(3.15)}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} \Sigma'_{Xij} = D[X']$$

Seega $v^T \Sigma_X v \geq 0$ ning kuna v oli valitud suvaliselt siis on Σ_X mittenegatiivselt määratud. Kuna n ja $X_1, X_2, \dots X_n$ olid valitud suvaliselt, siis on iga kovariatsioonimaatriks mittenegatiivselt määratud.
Sellega oleme andund reliaablusele matemaatilise definitsiooni ning mõtestanud lahti vajaminevad taustteadmised reliaabluse hindamiseks. Järgnevalt tuletame mõned alumised tõkked reliaablusele.

4 Alumised tõkked reliaablusele

Järgnevalt vaatleme viite alumist tõket reliaablusele, millest neli - $\lambda_1, \lambda_2, \lambda_3, \lambda_4$; pakkus aastal 1945 välja Louis Guttman [7]. Viienda vaadeldava tõkke pakkusid välja Jackson ja Anguwamba oma 1977 aasta artiklis [9]. Kõik tõkked tuletame tuginedes Jacksoni ja Anguwamba artiklile.
Olgu meil vastajate kogum I, üle mille me vaatleme k küsimusest koosnev küsimustik, mille vaadeldav tulemus on X, tegelik tulemus on T ja mõõtmisviga üle on E. Nummerdame küsimused arvudega 1 kuni k. Vaatleme igat 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 $\Sigma_X, \Sigma_T, \Sigma_E \in Mat_k(\mathbb{R})$. Nagu eelmises peatükkis, tähistame ka siin maatriksi Σ_X elemente kui x_{ij}, maatriksi Σ_T elemente kui t_{ij} ning maatriksi Σ_E peadiagonaali elemente θ_i. Edaspidi vaatleme selles peatükkis küsimustiku, mille vaadeldav tulemus on X, reliaabluse alumisi tõkkeid.

4.1 Alumine tõke λ_1

Vaatleme kovariatsioonimaatriksit Σ_T. Eeldame vastuväiteliselt, et maatriksi Σ_T peadiagonaalil leidub element, mis on negatiivne. Olgu selleks elemendiiks 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,

$$v^T \Sigma_T v = \begin{pmatrix} t_{i1} & t_{i2} & \dots & t_{ik} \end{pmatrix} v = t_{ii}.$$

Seega ei ole maatriks enam positiivselt määratud, mis on eelmises peatükkis läbiviidud arutelu põhjal vastuolus sellega, et Σ_T on kovariatsioonimaatriks .
Seega iga i 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 $\sum\limits_{i=1}^{k} \theta_i \leq \sum\limits_{i=1}^{k} x_{ii}$. Võttes arvesse võrdust (3.13), oleme saanud küsimustiku, mille vaadeldav tulemus on X, reliaablusele alumise tõkke, mida tähistame kui λ_1, kus

$$\lambda_1 = 1 - \frac{\sum_i x_i}{\sum_i \sum_j x_{ij}} \leq 1 - \frac{\sum_i \theta_i}{\sum_i \sum_j x_{ij}} = 1 - \frac{D\left[E\right]}{D\left[X\right]}. \tag{4.1}$$

4.2 Alumine tõke λ_2

Enne järgmise alumise tõke 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-negatiivlet määratud, siis peab kehtima

$$0 \leq v^T C v = \begin{pmatrix} \mathbf{0}^T & 1 & \mathbf{0}'^T & \alpha & \mathbf{0}''^T \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \\ \mathbf{0}' \\ \alpha \\ \mathbf{0}'' \end{pmatrix}, \tag{4.2}$$

kus $\mathbf{0}, \mathbf{0}', \mathbf{0}''$ on null vektorid pikkustega vastavalt $i-1, j-i-1, n-j$.
Arvutame avaldise 4.2 väärtuse:

$$v^TCv = \begin{pmatrix} \mathbf{0}^T & 1 & \mathbf{0}'^T & \alpha & \mathbf{0}''^T \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \\ \mathbf{0}' \\ \alpha \\ \mathbf{0}'' \end{pmatrix} =$$

$$= \begin{pmatrix} \mathbf{0}^T & 1 & \mathbf{0}'^T & \alpha & \mathbf{0}''^T \end{pmatrix} \begin{pmatrix} c_{1i} + \alpha c_{1j} \\ \vdots \\ a_{ni} + \alpha c_{nj} \end{pmatrix} = c_{ii} + \alpha c_{ij} + \alpha c_{ji} + \alpha^2 c_{jj} = c_{ii} + 2\alpha c_{ij} + \alpha^2 c_{jj}.$$

Seega saame, et $0 \leq c_{ii} + 2\alpha c_i j + c_{jj}$. See, et saadud ruutvõrratus on kehtib on samaväärne sellega, et selle 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 õppikust. [17].

Paneme tähele, et kehtib järgvev võrdus

$$\left(\sum_{i=1}^{k} x_{ii} \right)^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} x_{ii} x_{jj} = \sum_{i=1}^{k} x_{ii}^2 + \sum_{i=1}^{k} \left(\sum_{j=1}^{i-1} x_{ii} x_{jj} + \sum_{m=i+1}^{k} x_{ii} x_{mm} \right).$$

Järelikult kehtib ka

$$2(k-1) \left(\sum_{i=1}^{k} x_{ii} \right)^2 = 2(k-1) \sum_{i=1}^{k} x_{ii}^2 + 2(k-1) \sum_{i=1}^{k} \left(\sum_{j=1}^{i-1} x_{ii} x_{jj} + \sum_{m=i+1}^{k} x_{ii} x_{mm} \right).$$

Liidetavad teistmoodi grupeerides saame, et

$$2(k-1) \sum_{i=1}^{k} x_{ii}^2 + 2(k-1) \sum_{i=1}^{k} \left(\sum_{j=1}^{i-1} x_{ii} x_{jj} + \sum_{m=i+1}^{k} x_{ii} x_{mm} \right) =$$

$$= \left(x_{11}^2 - 2x_{11}x_{22} + x_{22}^2\right) + \left(x_{11}^2 - 2x_{11}x_{33} + x_{33}^2\right) + \cdots + \left(x_{11}^2 - 2x_{11}x_{kk} + x_{kk}^2\right) +$$

$$+ \left(x_{22}^2 - 2x_{22}x_{11} + x_{11}^2\right) + + \left(x_{22}^2 - 2x_{22}x_{33} + x_{33}^2\right) + \cdots + \left(x_{22}^2 - 2x_{22}x_{kk} + x_{kk}^2\right) +$$

$$+ \cdots +$$

$$+ \left(x_{kk}^2 - 2x_{kk}x_{11} + x_{11}^2\right) + \left(x_{kk}^2 - 2x_{kk}x_{11} + x_{11}^2\right) + \cdots + \left(x_{kk}^2 - 2x_{kk}x_{k-1k-1} + x_{k-1k-1}^2\right) +$$

$$+ 2k \sum_{i=1}^{k} \left(\sum_{j=1}^{i-1} x_{ii} x_{jj} + \sum_{m=i+1}^{k} x_{ii} x_{mm} \right) =$$

$$= \sum_{i=1}^{k} \left(\sum_{j=1}^{i-1} (x_{ii} - x_{jj})^2 + \sum_{m=i+1}^{k} (x_{ii} - x_{mm})^2 \right) + 2k \sum_{i=1}^{k} \left(\sum_{j=1}^{i-1} x_{ii} x_{jj} + \sum_{m=i+1}^{k} x_{ii} x_{mm} \right)$$

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

$$\sum_i \theta_i = \sum_i x_{ii} - \sum_i t_{ii}$$

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

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

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

$$\sum_i t_{ii} = \sqrt{\frac{\sum\limits_{i=1}^{k}\left(\sum\limits_{j=1}^{i-1}(t_{ii}-t_{jj})^2 + \sum\limits_{m=i+1}^{k}(t_{ii}-t_{mm})^2\right) + 2k\sum\limits_{i=1}^{k}\left(\sum\limits_{j=1}^{i-1} t_{ii}t_{jj} + \sum\limits_{m=i+1}^{k} t_{ii}t_{mm}\right)}{2\,(k-1)}} \geq$$

$$\geq \sqrt{\frac{k\sum\limits_{i=1}^{k}\left(\sum\limits_{j=1}^{i-1} t_{ii}t_{jj} + \sum\limits_{m=i+1}^{k} t_{ii}t_{mm}\right)}{(k-1)}}$$

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

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

Selle taustal saame defineerida alumise tõkke reliaablusele, mida tähistame λ_2:

$$\lambda_2 = 1 - \frac{\sum_i x_{ii} - \sqrt{\frac{k\sum\limits_{i=1}^{k}\left(\sum\limits_{j=1}^{i-1} x_{ij}^2 + \sum\limits_{m=i+1}^{k} x_{im}^2\right)}{(k-1)}}}{\sum_i \sum_j x_{ij}} \leq 1 - \frac{\sum_i x_i - \sum_i t_i}{\sum_i \sum_j x_{ij}} = 1 - \frac{\sum_i \theta_i}{\sum_i \sum_j x_{ij}} = 1 - \frac{D\,[E]}{D\,[X]}.$$

4.3 Alumine tõke $\lambda 3$

- - - SIIA MAANI REFAKTOREERITUD! EDASI OMAL VASTUTUSEL! - - - -

Vaatleme jällegi kovariatsioonimaatriksit Σ_T. Olgu meil n-vektor v, mille elment kohal i on 1 ja element kohal j on -1. Kuna kovariatsioonimaatriks peab olema mittenegatiivselt määratud siis peab kehtima

$$0 \leq v^T \Sigma_T v = v^T \begin{pmatrix} t_{1i} - t_{1j} \\ t_{2i} - t_{2j} \\ \vdots \\ t_{ni} - tnj \end{pmatrix} = t_i + t_j - 2t_{ij},$$

võttes arvesse seda, et kui $i \neq j$, siis $x_{ij} = t_{ij}$ saame, et $t_i + t_j \geq 2x_{ij}$.
Teame, et võimalusi valida i ja j nii, et $i \neq j$ on $n\,(n-1)$. Summeerides kõik võimalikud võrratused saame, et

$$2\,(n-1) \sum_i t_i \geq 2 \sum_i \sum_{\substack{j \\ i \neq j}} x_{ij},$$

mis on samaväärne sellega, et

$$\sum_i t_i \geq \frac{1}{n-1} \sum_i \sum_{\substack{j \\ i \neq j}} x_{ij}.$$

Niisiis, oleme saanud järgmise alumise tõkke reliaablusele,

$$\lambda_3 = 1 - \frac{\sum_i x_i - \frac{1}{n-1} \sum_i \sum_{\substack{j \\ i \neq j}} x_{ij}}{\sum_i \sum_j x_{ij}} \leq 1 - \frac{\sum_i x_i - \sum_i t_i}{\sum_i \sum_j x_{ij}} = 1 - \frac{\sum_i \theta_i}{\sum_i \sum_j x_{ij}} = 1 - \frac{D\,[E]}{D\,[X]}.$$

Tõkke lihtsustamisel algebraliste samaväärsustega saame, et

$$\lambda_3 = 1 - \frac{\sum_i x_i - \frac{1}{n-1} \sum_i \sum_{\substack{j \\ i \neq j}} x_{ij}}{\sum_i \sum_j x_{ij}} = \frac{\sum_i \sum_j x_{ij} - \sum_i x_i + \frac{1}{n-1} \sum_i \sum_{\substack{j \\ i \neq j}} x_{ij}}{\sum_i \sum_j x_{ij}} =$$

$$= \frac{\sum_i \sum_{\substack{j \\ i \neq j}} x_{ij} + \frac{1}{n-1} \sum_i \sum_{\substack{j \\ i \neq j}} x_{ij}}{\sum_i \sum_j x_{ij}} = \frac{n}{n-1} \left(\frac{\sum_i \sum_{\substack{j \\ i \neq j}} x_{ij}}{\sum_i \sum_j x_{ij}} \right) = \frac{n}{n-1} \left(1 - \frac{\sum_i x_i}{\sum_i \sum_j x_{ij}} \right).$$

Tuletades meelde tõkke λ_1 definitsiooni on ilmne, et

$$\lambda_3 = \frac{n\ \lambda_1}{n-1}. \tag{4.3}$$

4.4 Alumine tõke λ_4

Nagu eelmise kahe tõkku puhul, vaatleme kovariatsioonimaatriksit Σ_T. Olgu meil vektor v selline, et $|v_i| = 1$ iga $i \leq n, i \in \mathbb{N}$ korral. Teame, et kehtib

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

seega arvestades kovariatsioonimaatriksi Σ_E omadusi saame võrduse (3.11) põhjal

$$\sum_i \theta_i = \sum_i v_i^2 \theta_i = v^T \Sigma_E v \leq v^T \Sigma_X v.$$

Seega, iga võimaliku v korral saame alumise tõkke reliaablusele. Tõepoolest, (3.13) põhjal iga v korral,

$$1 - \frac{v^T \Sigma_X v}{\sum_i \sum_j x_{ij}} \leq 1 - \frac{\sum_i \theta_i}{\sum_i \sum_j x_{ij}} = 1 - \frac{D[E]}{D[X]}$$

Defineerime alumise tõkke λ_4 kui maksimumi kõikidest sellistest tõkketest, seega

$$\lambda_4 = \max_v \left(1 - \frac{v^T \Sigma_X v}{\sum_i \sum_j x_{ij}} \right).$$

4.5 Alumine tõke glb

Me teame, et kehtib (3.11) ning teame, et Σ_X, Σ_T ja Σ_E on kovariatsioonimaatriksid, seega ka mitte-negatiivselt määratud.
Paneme tähele, et tänu Σ_E struktuurile on tema mittenegatiivsus samaväärne sellega, et $0 \leq \theta_i$ iga i korral Tõestus? Iseenesest väga lihtne.
Üritame nüüd leida vea suurimat võimaliku dispersiooni, mis on (3.10) põhjal samaväärne summaga $\sum_i \theta_i$, arvestades eelnevalt kirjeldatud piiranguid. Saame selle panna kirja järgneva optimiseerimisülesandena:

$$\begin{aligned} max \sum_i \theta_i \\ \theta_i > 0 \\ \Sigma_X = \Sigma_T + \Sigma_E \\ \Sigma_T \succeq 0, \end{aligned} \tag{4.4}$$

kus kirjapilt $\Sigma_T \succeq 0$ tähistab tingimust "Σ_T on mittenegatiivselt määratud".
Olgu z eelneva optimiseerimisülesande lahend mingi fikseeritud vaadelud tulemuste kovariatsioonimaatrikis Σ_X korral. Tähistame tähekombinatsiooniga glb, mis

tuleneb fraasist *greatest lower bound*, z abil saadud alumise tõkku reliaablusele ehk

$$glb = 1 - \frac{z}{\sum_i \sum_j x_{ij}}.$$

Panemet 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 väiksem võrdne tõkega glb. Seega on glb suurim võimalik alumine tõke.

4.6 Cronbachi alfa

Cronbachi alfa on levinud sisemist reliaablust iseloomustav näitaja, mis on defineeritud järgnevalt:

$$(\frac{k}{k-1})(1 - \frac{\sum \sigma_i^2}{\sigma_t^2}) \tag{4.5}$$

kus k tähistab küsimuste arvu, σ_i on standardviga ühe küsimuse piires ja σ_t on standardviga üle testi kogutulemuste. [2, 396]
Siia selgitav näidis tabel/joonis vajalik?

Miks ma taandan selles töös sisemise järjekindluse mõõtmise Cronbachi alfale?

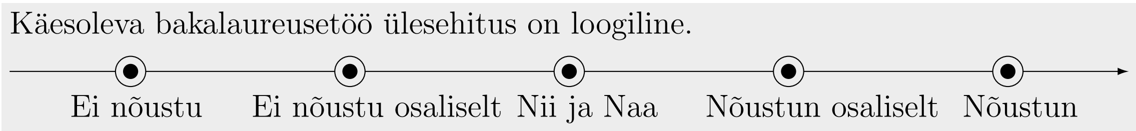
Kuna Cronbach'i alfal on mitmeid puuduseid, millest mõningad on toodud ära David Streineri artiklis [16, 101-102], ning isegi mõõdiku autor Lee Cronbach soovitab oma 1951. aastal välja töötatud mõõdiku asemel kasutatada alternatiivseid mõõdikuid [2], on oluline vastata alapealkirjas püstitatud küsimusele. Kaks peamist põhjust on järgnevad:

I Cronbachi alfa on hetkel kõige kasutatavam mõõdik sisemise reliaabluse mõõtmiseks psühho-meetriliste testide juures. Seetõttu on väljapakutud meetod testide tulemuste analüüsijatele tuttav ning seega loodetavasti lihtsamini kasutatav.

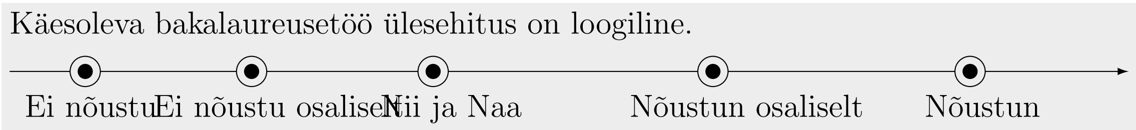
II Cronbachi alfa leidmine on arvutuslikult küllaltki lihtne. Kuna käesoleva töö peamine eesmärk on selgitada, kas väljapakutud lähenemine on mõistlik, siis ei ole paremematest mõõdikutest potentsiaalselt saadav kasu piisavalt suur, et tasa teha ülesande lahendamise raskemaks muutumist. Juhul, kui töö tulemus on aga positiivne, tuleks uurida ka teiste mõõdikute võimalikku kasutamist.

5 Ülesande püstitus

Selle töö raames loodame välja pakkuda viisi, kuidas paigutada intervallskaalal erinevaid vastuseid Likerti skaalalt. Eesmärk on leida parem paigutus kui vaikimisi meetod, kus eeldatakse, et erinevate hinnagute vahekaugused on samad. Vaatleme kahte näidet:



Joonis 6: Näide, kuidas hinnangud skaalal hetkel vaikimisi meetodit kasutades paigutuvad



Joonis 7: Näide hinnangute alternatiivsest paiknemisest skaalal

Kui me tahame välja pakkuda alternatiive vaikimisi hinnangule, peab meil olema põhjendus, miks välja pakutud lahendus kirjeldab reaalsust täpsemalt. Pakume välja järgmise lahenduse: üritame leida hinnangute paiknemist skaalal nii, et küsitluse sisemine reliaablus oleks võimalikult suur. Piirame ennast sellega, et hinnangute esialgne järjestus ei tohi muutuda. Sisemise järjekindluse maksimeerimise taandame antud töö käigus Cronbachi alfa maksimeerimisele.[4]

6 Ülesande matemaatiline püstitus

Olgu meil küsimustik n-väitega, kus iga näite kohta palutakse hinnangut 5-palli Likerti skaalal. Hinnanguid võime vaadelda kui juhuslikke suurusi $K_1, K_2, ..., K_n, K_i w$ mille muutumispiirkonnaks on hulk $\{1, 2, 3, 4, 5\}$. Toome sisse ka tähistused $p_{i\alpha}, i \in \{1, 2, ..., n\}, \alpha \in \{1, 2, 3, 4, 5\}$, kus $p_{i\alpha}$ tähistab olemasolevate andmete põhjal antud hinnangut tõenäosusele, et küsimusele K_i anti hinnang α.
Tuletame meelde Cronbachi alfa definitsiooni:

$$(\frac{k}{k-1})(1 - \frac{\sum \sigma_i^2}{\sigma_t^2})$$

[4]Põhjendus, miks selline taandamine on tehtud, on ära toodud peatükis **??**.

Paneme tähele, et toodud eelduste põhjal avalduks vaadeldava küsimustiku korral alfa järgnevalt:

$$(\frac{k}{k-1})(1-\frac{\sum \sigma_i^2}{\sigma_t^2}) = \frac{n}{n-1}\left(1-\frac{\sum\limits_{i=0}^{n} D(K_i)}{D(\sum\limits_{i=0}^{n} K_i)}\right)$$

Võttes arvesse, et $D(\sum\limits_{i=0}^{n} K_i) = \sum\limits_{i=0}^{n}\sum\limits_{j=0}^{n} COV(K_i, K_j)$ saame eelneva kirjutada järgnevalt:

$$(\frac{k}{k-1})(1-\frac{\sum \sigma_i^2}{\sigma_t^2}) = \frac{n}{n-1}\left(1-\frac{\sum\limits_{i=0}^{n} D(K_i)}{\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{n} COV(K_i, K_j)}\right)$$

Vastavalt ülesande püstituses toodud ideele soovime leida sellise viisi hinnangute tõlgendamiseks, et Cronbachi alfa oleks maksimaalne. Kujutame juhuslikud suurused $K_1, K_2, ..., K_n$ juhuslikeks suurusteks $L_1, L_2, L_2, ..., L_n$, kusjuures juhusliku suuruse L_i määramispiirkond ühtib suuruse K_i määramispiirkonnaga ning L_i saab väärtusi hulgast $\Lambda_i = \{\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, \lambda_{4i}, \lambda_{5i}\}$ $i \in 1, 2, ...n$. Lisaks kehtigu järgnevad kitsendused:

$$\lambda_{i1} < \lambda_{i2} < \lambda_{i3} < \lambda_{i4} < \lambda_{i5}$$

Siin peavad ikka ranged võrratused olema ,(0,0,0,0,1) ei kõlba.

$$K_i = 1 \implies L_i = \lambda_{i1}, K_i = 2 \implies L_i = \lambda_{i2}, \cdots, K_i = 5 \implies L_i = \lambda_{i5}$$

Paneme tähele, et Cronbachi alfa arvutamine, seega ka maksimeerimine, on endiselt keeruline. Siinkohal märgime, et hinnangute tõlgenduste juures ei huvita meid mitte absoluutne, vaid suhteline paigutus. Niisiis võime juhuslikke suurusi piirata järgnevate kitsendustega:

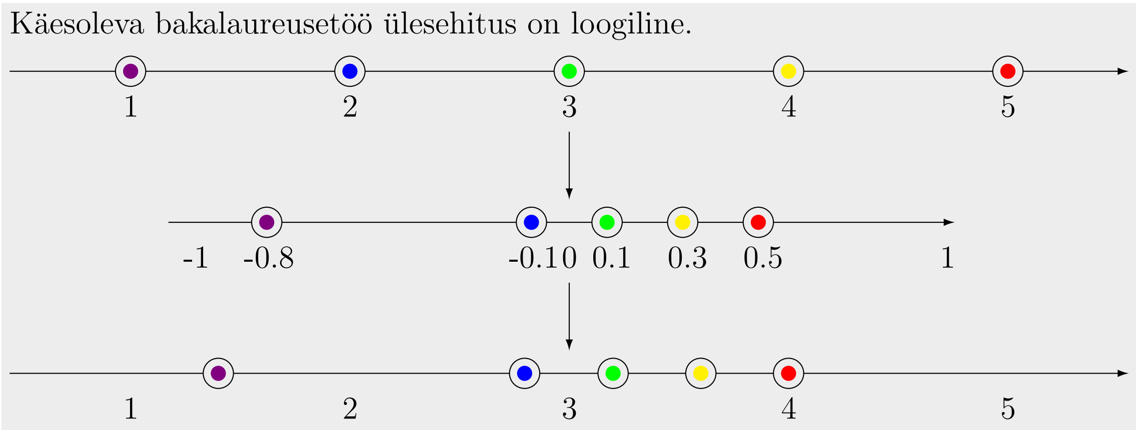
$$E(L_i) = p_{i1} * \lambda_{i1} + p_{i2} * \lambda_{i2} + p_{i3} * \lambda_{i3} + p_{i4} * \lambda_{i4} + p_{i5} * \lambda_{i5} = 0 \tag{6.1}$$

$$D(L_i) = p_{i1} * (\lambda_{i1})^2 + p_{i2} * (\lambda_{i2})^2 + p_{i3} * (\lambda_{i3})^2 + p_{i4} * (\lambda_{i4})^2 + p_{i5} * (\lambda_{i5})^2 = 1 \tag{6.2}$$

Paneme tähele, et sellisel skaalal olevaid küsimusi sisaldava testi Cronbachi alfa esitub lihtsamal kujul:

$$\alpha = \frac{n}{n-1}\left(1 - \frac{n}{\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{n} COV(L_i, L_j)}\right) \tag{6.3}$$

Siit näeme, et sellisel juhul taandub Cronbachi alfa maksimeerimine avaldise $\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{n} COV(L_i, L_j)$ maksimeerimisele. Järgev on täiesti trivaalne mapping Y = (X+1)*2 +1 vms. Põhjus miks ta siia sisse tõin oli soov rõhutada absoluutse skaala ebaolulisust. Võib välja jääda.Kuna uuritavaid teste vaadeldakse tavaliselt 5-palli skaalal, siis pärast sobivate kordajate leidmist võime teha veel ühe teisenduse, mis viib vahemikkust $(-1, 1)$ vahemikku $(0, 5)$, säilitades hinnangute (mida oleme tähistanud λ-dega) omavaheliste kauguste suhte. normeerime
Kogu eelnevalt tehtud illustreerib järgmine joonis, kus autor on püüdnud selgitada ühe juhusliku suuruse määramispiirkonna muutuse läbi eelnevalt kirjeldatud protsesside.



Joonis 8: Illustratsioon sellest, kuidas suhestub hulk $ran(K_i)$ hulka $ran(L_i)$ ning see omakorda hulka, mis tekib peale teisendust hulgast $ran(L_i)$ 5-palli skaalale, säilitades hinnagute vahelised kaugused

Olgu meil tõenäosuste maatriks P:

$$P = \begin{pmatrix} p_{(11)(11)} & p_{(11)(12)} & p_{(11)(13)} & p_{(11)(14)} & p_{(11)(15)} & p_{(11)(21)} & \cdots & p_{(11)(n5)} \\ p_{(12)(11)} & p_{(12)(12)} & p_{(12)(13)} & p_{(12)(14)} & p_{(12)(15)} & p_{(12)(21)} & \cdots & p_{(12)(n5)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{(n5)(11)} & p_{(n5)(12)} & p_{(n5)(13)} & p_{(n5)(14)} & p_{(n5)(15)} & p_{(n5)(21)} & \cdots & p_{(n5)(n5)} \end{pmatrix}$$

kus $p_{(i\alpha)(j\beta)}, i,j \in \{1,2,...,n\}, \alpha,\beta \in \{1,2,3,4,5\}$ tähistab tõenäosust, et küsimusele K_i anti vastus α ja küsimusele K_j anti vastus β.
Paneme tähele, et kuna $p_{(i\alpha)(i\alpha)} = p_{i\alpha}$, siis avaldub eelnev maatriks ka järgnevalt:

$$P = \begin{pmatrix} p_{i\alpha} & p_{(11)(12)} & p_{(11)(13)} & p_{(11)(14)} & p_{(11)(15)} & p_{(11)(21)} & \cdots & p_{(11)(n5)} \\ p_{(12)(11)} & p_{i\alpha} & p_{(12)(13)} & p_{(12)(14)} & p_{(12)(15)} & p_{(12)(21)} & \cdots & p_{(12)(n5)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{(n5)(11)} & p_{(n5)(12)} & p_{(n5)(13)} & p_{(n5)(14)} & p_{(n5)(15)} & p_{(n5)(21)} & \cdots & p_{i\alpha} \end{pmatrix}$$

Defineerime vektori x:

$$x = (\lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{15}, \lambda_{21}, \lambda_{22}, \lambda_{23}, \lambda_{24}, \lambda_{25}, \cdots, \lambda_{n1}, \lambda_{n2}, \lambda_{n3}, \lambda_{n4}, \lambda_{n5})$$

Siis $xPx^T = \sum_{i=1}^{n}\sum_{j=1}^{n} COV(L_i, L_j)$. Veendume selles:

$$xPx^T = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{n5} \end{pmatrix} \begin{pmatrix} p_{(11)(11)} & p_{(11)(12)} & \cdots & p_{(11)(n5)} \\ p_{(12)(11)} & p_{(12)(12)} & \cdots & p_{(12)(n5)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{(n5)(11)} & p_{(n5)(12)} & \cdots & p_{(n5)(n5)} \end{pmatrix} \begin{pmatrix} \lambda_{11} \\ \lambda_{12} \\ \vdots \\ \lambda_{n5} \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{j=1}^{n}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(11)} & \sum_{j=1}^{n}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(12)} & \cdots & \sum_{j=1}^{n}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(n5)} \end{pmatrix} \begin{pmatrix} \lambda_{11} \\ \lambda_{12} \\ \vdots \\ \lambda_{n5} \end{pmatrix} =$$

$$= \sum_{j=1}^{n}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(11)} + \sum_{j=1}^{n}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(12)} + \cdots + \sum_{j=1}^{n}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(n5)} =$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{5}\sum_{j=1}^{n}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(ik)} \lambda_{ik} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{5}\sum_{l=1}^{5} \lambda_{jl} p_{(jl)(ik)} \lambda_{ik} = \sum_{i=1}^{n}\sum_{j=1}^{n} E(L_i L_j) =$$

$$\underset{(1)}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} E(L_i L_j) - E(L_i)E(L_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} COV(L_i, L_j)$$

Eelneva põhjal piisab meile Cronbachi alfa leidmiseks ruutvõrrandi xPx^T maksimeerimisest, arvestades eelnevalt äratoodud piiranguid L_i keskväärtusele ja dispersioonile.. Selle põhjal saame püstitada *Quadratically constrained quadratic programm(QCQP)*'i Mingi pädev eestikeelne termin selle kohta ? tüüpi optimeerimisprobleemi, mille lahendus annab meile otsitavad tõlgendused. Teeme seda:

$$
\begin{aligned}
&min\ x^T(-P)x \\
&R_i^T x = 0, i \in 1,2,...,n \\
R_i = (&\underbrace{0,0,...0,0}_{(i-1)*5} p_i a, p_i b, p_i c, p_i d, p_i e, \underbrace{0,0,...,0,0}_{(n-i)*5}) \\
&x^T P_i x = 1, i \in \{1,2,...,n\},\ P_i = \begin{pmatrix} p_{ja} & 0 & \cdots & 0 \\ 0 & p_{jb} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{je} \end{pmatrix} \\
&p_j\alpha = \begin{cases} 0 & j \neq i \\ p_i\alpha & i = j \end{cases}, \alpha \in \{a,b,c,d,e\}
\end{aligned}
\tag{6.4}
$$

7 Ülesande tüübi keerekuse analüüs

Järgnevas analüüsime ruutkitsendustega ruuplaneerimisülesannete keerukust üldisemalt. Kõigepealt defineerime üldkuju, selleks on,

$$
\begin{aligned}
&min\ \frac{1}{2}x^T P_0 x + q_0^T x \\
&\frac{1}{2}x^T P_i x + q_i^T x + r_i \leq 0, i \in \{1,...,m\}, \\
&Ax = b, P_0, P_i \in R^{n\times n}, x, q_0, q_i \in R^n, \text{kus x on optimeeritav muutuja.}
\end{aligned}
\tag{7.1}
$$

Paneme tähele, et kitsendused $x_i(x_i - 1) \leq 0$, $x_i(1 - x_i) \leq 0$ rahuldavad probleemi (7.1) kitsendustele pandud tingimusi. Tõepoolest, toodud kitsendused saab viia kujule

$$
\begin{aligned}
&x^T P_i x + (-q_i^T)x \leq 0, \\
&x^T(-P_i)x + q_i^T x \leq 0, P_i = \begin{cases} p_i i = 1 \\ p_j k = 0, j, k \neq i \end{cases}, q_i = \begin{cases} q_i = 1 \\ q_j = 0, j \neq i \end{cases}.
\end{aligned}
$$

Eelnevad kitsendused on samaväärsed kitsendusega $x_i(x_i - 1) = 0$, mis on samaväärne kitsendusega $x_i \in \{0, 1\}$. Niisiis saab kahendtäisarv-programmi (binary integer program, 0-1 integer program) esitada eelnevalt kirjeldatud kujul.

7.1 Kahendtäisarv programm

Kahendtäisarv-programm on optimeerimisprobleem, kus kõik muutujad peavad saama väärtusi hulgast $\{0, 1\}$. Kahendtäisarv-programmi kanooniline kuju on

$$\begin{gathered} max\ c^T x \\ Ax \leq b, \\ x_i \in \{0, 1\}, \\ a \in A, b_0 \in b, c_0 \in c \implies a, b_0, c_0 \in \mathbb{Z} \end{gathered} \tag{7.2}$$

7.2 NP-Hard probleemid

8 Ülesande analüüsiks vajalik taust

8.1 Kumerus

Definitsioon 8.1. Olgu

9 Lahenduse idee

Kas probleem on SDP tüüpi või SDP? Miidlale kirjutan vast nendes küsimustes. Mingi sissejuhatav jutt QCQP ja SDP kohta, kas algusessse taust peatükki või pigem teksti sisse ? Püstitatud optimeerimis probleemi lahendamiseks teisendame probleemi *semidefinite programming(SDP)* tüüpi optimeerimisülesandeks ning seejärel lahendame saadud ülesande.

10 Teisendus $QCQP$ tüüpi ülesandelt SDP tüüpi ülesandele

Esiteks teisendame saadud $QCQP$ standartkujule. Siia vaja selgitusi, mis ja kuidas

$$\begin{gathered} min\ \ -\theta \\ x^T P x - \theta \leq 0 \\ R_i^T x \leq 0, i \in 1, 2, ..., n \\ -R_i^T x \leq 0, i \in 1, 2, ..., n \\ R_i = (\underbrace{0, 0, ...0, 0}_{(i-1)*5} p_i a, p_i b, p_i c, p_i d, p_i e, \underbrace{0, 0, ..., 0, 0}_{(n-i)*5}) \end{gathered} \tag{10.1}$$

$$x^T P_i x - 1 \leq 0, i \in \{1, 2, ..., n\}$$
$$-(x^T P_i x - 1) \leq 0, i \in \{1, 2, ..., n\}$$
$$P_i = \begin{pmatrix} p_{ja} & 0 & \cdots & 0 \\ 0 & p_{jb} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{je} \end{pmatrix}$$
$$p_j\alpha = \begin{cases} 0 & j \neq i \\ p_i\alpha & i = j \end{cases}, \alpha \in \{a, b, c, d, e\}$$

[6, 116]
Paneme tähele, et kuna maatriksid P, R_i, P_i on sümmeetrilised ja mittenegatiivselt määratud, siis leiduvad maatrikisd Q, S_i, Q_i nii, et $P = QQ^T, R_i = S_iS_i^T, P_i = Q_iQ_i^T$. Nende maatriksite leidmiseks saab kasutada Cholesky lahutust. [17, 151] raamatus juhj positiivselt määratud maatriksi korral, mittenegatiivselt määratud maatriksi korral sellied maatriksid leiduvad, kuid ei ole alati üheselt määratud.
Olgu A mittenegatiivselt määratud maatriks, olgu $A = LL^T$.Märkame, et kehtib järgnev:

$$x^T Ax \leq b^T x + c \iff \begin{pmatrix} I_k & L^T x \\ x^T L & b^T x + c \end{pmatrix} \succeq 0 \tag{10.2}$$

[11, 31]
Seega saame anda oma optimeerimisprobleemile järgneva kuju:

$$\begin{gathered} min \ \ -\theta \\ \begin{pmatrix} I_k & Q^T x \\ x^T Q & \theta \end{pmatrix} \succeq 0 \\ R_i^T x \leq 0, i \in 1, 2, ..., n \\ -R_i^T x \leq 0, i \in 1, 2, ..., n \\ R_i = (\underbrace{0, 0, ...0, 0}_{(i-1)*5} p_ia, p_ib, p_ic, p_id, p_ie, \underbrace{0, 0, ..., 0, 0}_{(n-i)*5}) \\ \begin{pmatrix} I_k & Q_i^T x \\ t^T & 1 \end{pmatrix} \succeq 0 \\ \begin{pmatrix} I_k & Q_i^T x \\ -x^T Q_i & -1 \end{pmatrix} \succeq 0 \end{gathered} \tag{10.3}$$

11 Teek

Põhjendan siin ka lühidalt Pythoni eeliseid teegi realiseerimisel, alternatiividega (näiteks Java või MATLAB) võrreldes:

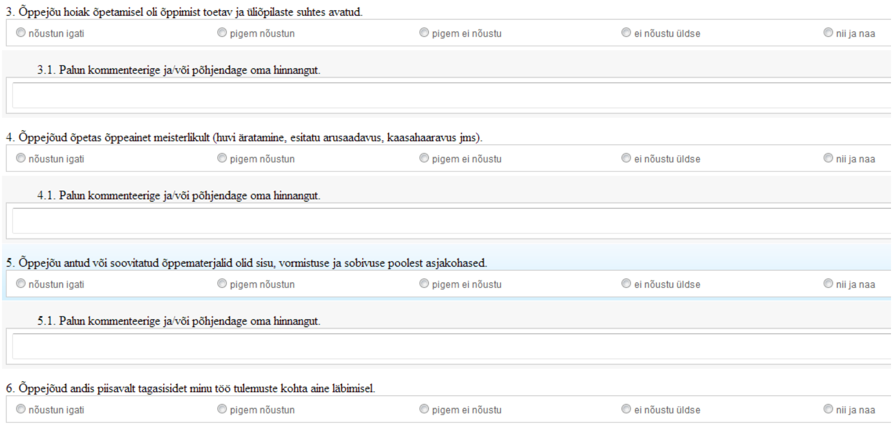
1. Python on hetkel keel, millega programeerimisalased õppingud tavaliselt algavad. Näiteks Tartu Ülikoolis õpetatakse programeerimist keeles Python nii matemaatika-informaatika teaduskonnale suunatud aines kui ka teistele huvilistele suunatud aines "Programeerimise alused", mida 2013/2014 aasta kevadel kuulas 150 inimest. Autor loodab, et nende seas on ka mõned psühholoogia tudengid.

2. Pythonil on läbi NumPy ja SciPy teekide väga hea arvutuslik tugi soovitud hinnangute implementeerimiseks.

3. Pythoni teek on väga lihte integreerida mõne veebiraamistikuga, mis võiks olla loodava teegi üks rakenduskohti.

4. Python on vabavara ning kõigile tasuta kasutatav.

12 Lisa


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.

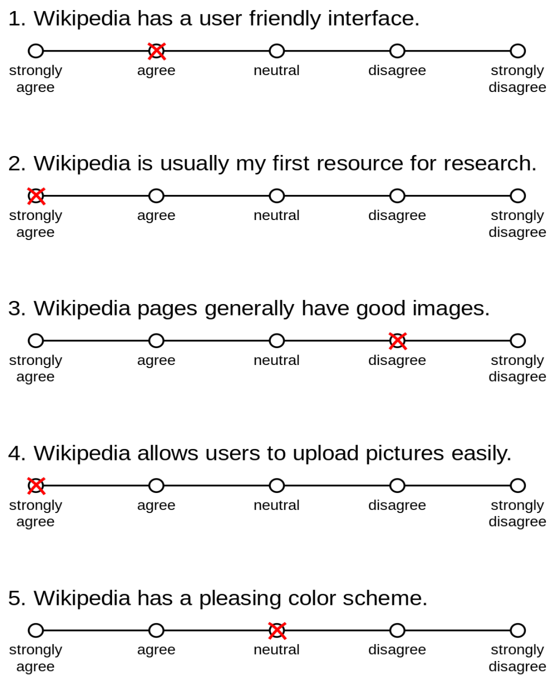
nõustun igati | pigem nõustun | pigem ei nõustu | ei nõustu üldse | nii ja naa

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



Joonis 10: Näide küsimustikust, kus on rakendatud Likerti skaalat; küsimused on paigutatud nende järjestikulisuse rõhutamiseks teljele [15]

Viited

[1] Lee J. Cronbach. Test "Reliability": Its Meaning and Determination. *Psychometrika*, 12(1):1–16, 1947.

[2] 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.

[3] Dato N. M. de Gruijter and Leo J. Th. van der Kamp. *Statistical Test Theory For Education And Psychology.* Number October. Taylor & Francis, 2007.

[4] N. Dilmen. Reliability and Validity, 2008.

[5] Diane R. Edmondson. Likert Scales: A History. http://faculty.quinnipiac.edu/charm/CHARM%20proceedings/CHARM%20article%20archive 2005.

[6] M. A. Epelman. Introduction to Semidefinite Programming (SDP). http://www-personal.umich.edu/~mepelman/teaching/IOE511/Handouts/511notes07-16.pdf, 2007.

[7] Louis Guttman. A Basis for Analyzing Test-Retest Reliability. *PSYCHOMETRIKA*, 1[(4):255–282, 1945.

[8] Robin K. Henson. Understanding Internal Consistency Reliability Estimates: A Conceptual Primer on Coefficient Alpha. *Measurement & Evaluation in Counseling & Development*, 34(3):177, 2001.

[9] 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.

[10] Susan Jamieson. Likert scales: how to (ab)use them. *Medical education*, 38(12):1217–8, December 2004.

[11] Monique Laurent and Frank Vallentin. Semidefinite Optimization. http://page.mi.fu-berlin.de/fmario/sdp/laurentv.pdf, 2012.

[12] F.M. Lord and M.R Novick. *Statistical theories of mental test scores.* 1968.

[13] Melvin R. Novick. The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1):1–18, February 1966.

[14] Klaas Sijtsma. On the use, the misuse, and the very limited usefulness of cronbach's alpha. *PSYCHOMETRIKA*, 74(1):107–120, 2009.

[15] Nicholas Smith. Example of Likert Scale. http://en.wikipedia.org/wiki/File:Example_Likert_Scale.svg.

[16] David L Streiner. Starting at the Beginning : An Introduction to Coefficient Alpha and Internal Consistency Starting at the Beginning : An Introduction to Coefficient Alpha and Internal Consistency. (February 2013):37–41, 2010.

[17] Ivar Tammeraid. *Lineaaralgebra rakendused*. TTÜ Kirjastus, Tallinn, 1999.

[18] UT. Õisi tagasiside. http://is.ut.ee.