

Tartu Ülikool

Loodus- ja täppisteaduste valdkond

Matemaatika ja statistika instituut

HEAOLU UURING SUVAMAAL

Projekt Valikuuringute teooria I kursuse raames

Autorid: Eesnimi1 Perenini1
Eesnimi2 Perenini2

Tartu 2022

1 Sissejuhatus

Suvamaal on viimasel ajal hakanud levima väited, et inimesed on õnnetud ning nende tervis on kehv. Nad on täheldanud, et elurahulolu on alla keskmise igal juhul, aga reaalsuses mingeid andmeid selle kohta ei olnud. Suvamaa otsustas palgata värskeid statistikuid, et nad koostaksid uuringu just selle probleemi raames. Valitsus andis neile vabad käed, kuidas teostada uuring ning kuidas analüüsida andmeid, aga nõudis kindlaid vastuseid.

Valitsus soovis kindlaid vastuseid, et arvestada sellega järgmise aasta eelarve planeerimises. Kui tõesti kuulujuttudel alust on, siis tuleb rohkem rõhku panna palkade tõstmisele, riigitoetustele, vaimse tervise spetsialistidele ning spordile, kuid asjata ei soovi valitsus raha kulutada asjadele, mis ei pruugi Suvamaa elanikke üldse aidata.

Statistikud koostasid uuringu selliselt, et nad küsisid inimestelt, kui palju nad nädalas tööd teevad, mis nende vanus on, mis nende hinnangud õnnelikkusele, tervisele ning rahulolule on. Viimasena uuriti ka, kas elanik on kuidagi seotud usuga, sest **elanike ja usklike** hinnangutel võib olla erinevus sees.

Valitsus kuulas plaani ära ning kiitis selle heaks.

Usklikud on ka elanikud.
Korrektsem lause oleks, "sest usklike hinnangud võivad erineda kõigi elanike hinnangutest"

2 Metoodika

Analüüsis kasutame kogusumma ning keskmise hinnangut. Kogusumma üldkogumis saame leida kujul:

$$t = \sum_U y_i \quad (1)$$

kus t on kogusumma, U on üldkogum ning y_i on tunnuse väärtus. Üldkogumi keskmine avaldub kujul:

$$\bar{Y} = \frac{t_y}{N} \quad (2)$$

kus t_y on mingi tunnuse kogusumma ning N on üldkogumi maht. Kasutame üldist hindamisteoreemi (edaspidi ÜHT), et valimi põhjal hinnata üldkogumi tunnuse kogusummat. ÜHT ütleb meile, et üldkogumi nihketa hinnang kogusummale on järgmine:

$$\hat{t} = \sum_s I_i \check{y}_i = \sum_s w_i y_i, \quad \text{kus} \quad \check{y}_i = \frac{y_i}{EI_i} \quad \text{ja} \quad w_i = \frac{I_i}{EI_i} \quad (3)$$

I_i on i -nda vaatluse kaasamiste arv. Keskmise hinnangu saame nüüd eelneva põhjal avaldada nii:

$$\hat{\bar{Y}} = \frac{\hat{t}}{N} \quad (4)$$

Hinnangu \hat{t} valikupõhine dispersioon on

$$V(\hat{t}) = \sum_{i=1}^N \sum_{j=1}^N c_{ij} \Delta_{ij} \check{y}_i \check{y}_j$$

c_{ij} on dispersiooni hinnangu tõestamise jaoks lisatud tundmatu konstant. Siin peaks olema valem (13) leheküljelt 24

Selle nihketa hinnanguks saame

$$\hat{V}(\hat{t}) = \sum_U \sum_U \check{\Delta}_{ij} w_i y_i w_j y_j \quad (6)$$

kus w_i ning w_j on valikukaalud ning $\Delta_{ij} = Cov(I_i, I_j)$. Suhtelise vea leiame eelmiste tulemuste abil:

$$\frac{\sqrt{\hat{V}(\hat{t}_y)}}{\hat{t}_y} \quad (7)$$

Tasub ka ära märkida 95% usaldusintervalli valemi kogusumma puhul

$$\hat{t} \pm 1.96 \sqrt{\hat{V}(\hat{t})} \quad (8)$$

Samuti ka suhte hinnangu valem

$$\hat{R} = \frac{\hat{t}_y}{\hat{t}_x} \quad (9)$$

Paneme kirja ka samad hinnangud osakogumi jaoks. Kogusumma osakogumis:

$$\hat{t}_d = \sum_U w_i y'_i, \quad (10)$$

y'_i on osakogumis asuva tunnuse väärtus. Keskmise saame avaldada kogusumma hinnangu järgi

$$\hat{Y}_d = \frac{1}{N_d} \hat{t}_d \quad (11)$$

Dispersiooni hinnangu valem jääb samaks, ainult kasutada uut tunnust y'_i . Suhtehinnang on nüüd kujul

$$\hat{R}_d = \frac{\sum_U w_i y'_i}{\sum_u w_i x'_i} \quad (12)$$

(Konspekt aines Valikuuringute Teooria I (lk 22-25, 29, 31, 33, 35-37). N. Lepik, I. Traat, K. Lumiste)

Üldised valemid on korrektsed, aga võiks lisada ka kasutatud disaini kogusumma ja disperisooni valemid.

3 Uuritavad tunnused ja valikudisain

3.1 Tunnused

Selleks, et saada aimu, missugune on rahulolu ja kuidas inimesed oma elu hindavad, oleme kasutusele võtnud järgnevad tunnused

1. wkhtot - Nädalas töötatud tundide arv (ületunnid kaasarvatud). **kaasa arvatud**

Sellega näeme, kui palju tunde on töötatud nädalas kogu rahvaarvu poolt. Seda saab hiljem uuritava rahvastiku poolt.

Rahvaarv on statistiline näitaja, rahvaarv ei tee tööd.

2. agea - Inimese vanus

Seda tunnust kasutame, et leida 50+ inimeste osakaalu hinnang.

3. happy - Inimese õnnelikkuse hinnang (1-10)

Sellega mõõdame valimi keskmist õnnelikkuse taset.

4. health - Inimese tervises seisundi hinnang (1-5)

Lisada ka skaalade selgitused, siis ei teki vigu hiljem tulemuste analüüsimisel

Mõõdame valimi keskmist tervises seisundi hinnangut.

5. stflife - rahulolu hinnang elule (1-10)

Kuidas hindab inimene enda rahulolu üldiselt?

6. rlgblg - Kas on mingisuguse uskumusega seotud? (1 - Jah, 0 - Ei)

Kasutame, et leida osakogum, kus inimesed on religiooniga seotud.

3.2 Valikudisain

Käesolevas analüüsis kasutame valimi võtmiseks üldkogumist lihtsat juhuslikku valikut, mis on tagasi-panekuta (edaspidi LJV TTA). LJV TTA valikudisain on fikseeritud mahuga.

Valimimahuks on määratud 200 objekti ning üldkogumi suurus on 1904, seega valim moodustab umbes 10.5% üldkogumist. Iga objekti kaasamistõenäosus on täpselt ühesugune ehk $\frac{n}{N}$, kus n on valimi suurus ning N on üldkogumi suurus.

Valimis me loeme usklikeks kõiki, kes on kuidagi mingisuguse religiooniga seotud on. Samuti nimetame vanemateks inimesteks kõiki, kes on 50+ aastat vana ning kasutame seda osakaalu leidmiseks.

Lauses 2x "on"

vanad

Eesti keeles on arvudes kümnendkoha eraldamiseks koma, inglise keeles punkt. Valim moodustab umbes 10,5% üldkogumist

4 Tulemuste leidmine

1. Kasutasime paketti survey ning leidsime hinnangu nädala töötundide kohta: 77312 tundi nädalas ning standardhälve tuli 2931. Suhteline viga on 3.79%. (Lisa 3. Kogusumma hinnang).
2. Järgmiseks leidsime hinnangud õnnelikkuse, tervise ja rahulolu keskmistele. (Lisa 4. Keskmiste hinnang)

Tunnus	Keskmise hinnang	Standardhälbe hinnang	Suhteline viga	95% intervall
Õnnelikkus	7.38	0.137	1.85%	(7.11, 7.65)
Tervis	2.49	0.061	2.47%	(2.37, 2.61)
Rahulolu	7.04	0.149	2.12%	(6.75, 7.33)

Tabel 1. Tunnuste keskmiste hinnangud.

3. Usklike inimeste hinnatavaks osakaaluks saime 24%, standardhälve 0.03 ning hinnangu suhteline viga on 3.03% (Lisa 5. Usklike osakaalu hinnang). Usklike osakogumis on nädala töötundide kogusumma hinnanguks 18050, standardhälbe hinnang on 2392.5 ning suhteline viga tuli 0.126 (Lisa 6. Usklike töötundide hinnang). Arvestades, et see hinnang on 24% kohta, siis kui teisendada kõikide elanike töötundide kogusumma hinnang samasuguse osakaalu peale, saame ligikaudu 18 554 hinnanguks. **Erineva suurusega osakogumite võrdlemiseks kasutatakse kogusumma asemel keskmist näitajat. Näiteks keskmine töötundide arv töötava inimese kohta**
4. Leiame hinnangud õnnelikkuse, tervise ja rahulolu keskmistele usklike osakogumis. (Lisa 7. Keskmiste hinnangud osakogumis)

Väärtused esitada ühesuguse täpsusega

Tunnus	Keskmise hinnang	Standardhälbe hinnang	Suhteline viga	95% intervall
Õnnelikkus	7.56	0.252	3.33%	(7.07, 8.05)
Tervis	2.6	0.117	4.5%	(2.37, 2.83)
Rahulolu	7.08	0.327	4.62%	(6.44, 7.72)

Tabel 2. Osakogumis keskmiste hinnangud. **Usklike osakogumis keskmiste hinnangud**

hinnangu standardhälve = standardviga

5. Vanemate inimeste hinnatav osakaal valimis on 50%, **standardhälve** on hinnanguliselt 0.035 ning suhteline viga 7.09% (Lisa 9. 50+ inimeste osakaal valimis). Usklike osakogumis on hinnanguliselt aga 68.75% vanemaid inimesi, **standardhälve** 0.067 ning suhteline viga tuleb 12.6%. (Lisa 8. 50+ inimeste osakaal usklike seas) **standardviga**

(Hinnangute leidmiseks kasutatud valemeid vaata lehekülgedelt 3 ja 4).

5 Analüüs

Sel hetkel, kui Suvamaal hakkasid kuulujutud üle jõu käima, said statistikud kõik oma andmed kogutud ning vajalikud analüüsid tehtud. Tulemused olid üllatuslikud.

Kuulujuttude järgi on Suvamaa elanikud õnnelikud. Teostatud uuringuga saadud andmetest andmestest sai aga välja lugeda, et õnnelikkuse keskmiseks hinnanguks on 7.38 10st, ehk inimesed on keskmisest õnnelikumad.

Järgmiseks uuriti välja, missugune on keskmine hinnang õnnelikkusele usklike kogukondades ning sooviti teada, kas ja kui palju see üldisest keskmisest erineb. Tuleb välja, et inimesed, kes puutuvad mingit moodi usuga kokku, hindavad oma õnnelikkust 7.56 punkti peale.

Kuna Suvamaa 95% usaldusintervall on (7.11, 7.65) ja usklike oma on (7.07, 8.05), ehk siis vastavalt on 95% tõenäosus, et nende keskmised jäävad sinna vahemikku, siis ei ole meil alust arvata, et nende keskmised hinnangud kuidagi erinevad. Seega võib öelda, et usklike õnnelikkuse hinnang ei erine kõigi omast. Tunnuse health skaala on vastupidine (1 - väga hea, 5 - väga halb). Seetõttu on toodud järelused vastupidised.

Tervise osas ei leitud aga nii häid tulemusi. Valimi järgi leitud hinnang keskmisele tervisele oli natukene alla keskmise, 2.49/5st. Usklikud inimesed aga hindasid on tervist parememaks, 2.6/5st. Statistikud arvasid, et sellega võib olla seotud 50+ inimeste osakaal, kuid tuli välja, et usklike osakogumis on 69.75% inimestest üle 50 aasta vanad, aga terves valimis on neid 50%.

Viimasena uuriti väidet, et Suvamaal ei olda eluga üldse rahul. Selleks uuriti elanikelt nende enda elu rahulolu hinnang ning arvutati välja riigi keskmine. Teostatud arvutused ütlevad aga pigem, et elanikud on üsnagi rahul enda eluga.

Leiti, et keskmine rahulolu hinnang Suvamaal on 7.04/10, mis viitab asjaolule, et tegemist on alusetute kuulujuttudega. Igaks juhuks vaadati ka eraldi usklike inimesi ning leiti, et nende keskmine tuleb 7.08/10. Seega, kuna keskmised oluliselt ei erine, siis võib öelda, et kõikide elanike ja usklike keskmine rahulolu hinnang on sama.

Kuna hinnanguid võis mõjutada ka töötundide arv, siis tasub ka neid vaadata. Elanike töötundide kogusumma hinnanguks tuli 77312. 24% sellest on 18 554 ning see annab meile võimaluse võrrelda töötundide kogusummat usklike omaga. Usklikel oli töötundide kogusumma hinnang 18050. Kuna erinevus märkimisväärne ei olnud, siis pole sellel ka erilist mõju hinnangutele.

Tänu statistikute analüüsile sai valitsus viimaks vajaliku lükke muutuseid sisse viia ning viimasel istungil jõuti üheskoos arusaamale, et nii vaimse kui ka füüsilise tervise fondi tuleb suurendada. Kuna valitsusel raha liiast pole ning nii õnnelikkuse tase kui ka rahulelu eluga oli elanike seas üle keskmise, siis otsustati muude hüvede jaoks seekord raha mitte juurde anda. rahulolu

sõnastus -
lauses kaks
erineva
tähtsusega
keskmist

6 Lisad

Lisa 1. Üldkogumi sisselugemine ning tunnuste valimine

```
library(readr)
library(dplyr)
library(sampling)
library(labelled)
library(survey)
setwd(dirname(rstudioapi::getActiveDocumentContext()[[2]]))
yldkogum <- readRDS("yldkogum_2022.Rdata")
yldkogum <- yldkogum %>%
  select(stflife, happy, health, agea, rlgblg, wkhtot)
yldkogum <- yldkogum %>%
  mutate(rlgblg = ifelse(rlgblg == 1, 1, 0), inimene = 1,
  vanema = ifelse(agea >= 50, 1, 0))
var_label(yldkogum$agea) <- NULL
```

Lisa 2. Valimi ning disaini moodustamine

```
valimimaht = 200
N = nrow(yldkogum)
valim <- yldkogum %>%
  mutate(valik = srswor(valimimaht, N))
valim <- valim %>% filter(valik == 1)
valim$kaal = N/valimimaht
disain.LJVTTA <- svydesign(ids = ~1, weights = ~kaal, data = valim)
```

Lisa 3. Nädala töötundide koguarvu hinnang

```
svytotal(~wkhtot, disain.LJVTTA, na.rm = T)
cv(svytotal(~wkhtot, disain.LJVTTA, na.rm = T))
```

Lisa 4. Keskmise õnnelikkuse, tervise ning rahulolu hinnang

```
svymean(~happy + health + stflife, disain.LJVTTA, na.rm = T)
cv(svymean(~happy + health + stflife, disain.LJVTTA, na.rm = T))
```


Lisa 5. Usklike osakaalu leidmine

```
svyratio(~rlgblg , ~inimene , disain.LJVTTA, na.rm=T)
cv(svyratio(~rlgblg , ~inimene , disain.LJVTTA, na.rm=T))
```

Lisa 6. Usklike nädala töötundide summa

```
svytotal(~wkhtot , subset(disain.LJVTTA,
                           rlgblg == 1), na.rm = T)
cv(svytotal(~wkhtot , subset(disain.LJVTTA,
                           rlgblg == 1), na.rm = T))
```

Lisa 7. Keskmise õnnelikkuse, tervise ning rahulolu hinnang usklike seas

```
svymean(~ happy + health + stflife ,
        subset(disain.LJVTTA, rlgblg == 1), na.rm = T)
cv(svymean(~ happy + health + stflife ,
          subset(disain.LJVTTA, rlgblg == 1), na.rm = T))
```

Lisa 8. 50+ inimeste osakaal usklike seas

```
svyratio(~vanema , ~inimene ,
        subset(disain.LJVTTA, rlgblg==1), na.rm = T)
cv(svyratio(~vanema , ~inimene ,
          subset(disain.LJVTTA, rlgblg==1), na.rm = T))
```

Lisa 9. 50+ inimeste osakaal valimis

```
svyratio(~vanema , ~inimene , disain.LJVTTA, na.rm=T)
cv(svyratio(~vanema , ~inimene , disain.LJVTTA, na.rm=T))
```