

STATISTIKA JA ANDMEANALÜÜS

Seminar 1

Kairi Osula

www.tlu.ee/~kairio

www.tlu.ee/~kairio/7410

kairio@tlu.ee

23.02.2018

Kursuse sisu

Empiiriline uuring. Andmete/tunnuste tüübid.

Kirjeldav statistika. Esmane andmete analüüs. Andmete graafiline kirjeldamine. Kirjeldavad arvnäitajad. Normaaljaotuse idee. Üldkogum ja valim. Statistiline üldistamine.

Parameetrilised ja mitteparameetrilised analüüsimeetodid.

Statistiliste hüpoteeside kontrollimine. Korrelatsioonanalüüs. Seosekordajate statistiline olulisus. Lineaarne regressioonanalüüs.

Eksam

Eksami eeldused:

- **rühmatöö nr. 1**, tähtaeg 9.03.2018
- **rühmatöö nr. 2**, tähtaeg 6.04.2018

Tähtajast hiljem esitatud eksami eeldused on ind. tööd. Esitada hiljemalt 1 nädal enne eksami kuupäeva. Kui eksamieeldus on arvestamata, eksamile tulla ei saa.

Eksam:

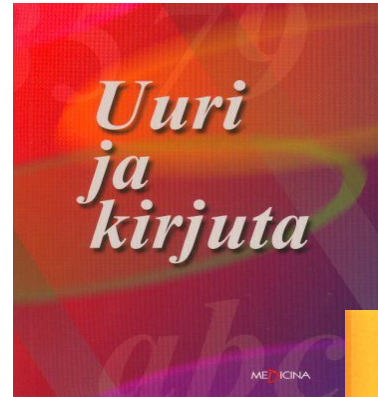
- (50%) kirjalik, avatud küsimustega test.
- (50%) praktiline andmeanalüüs (individuaalne töö)

„A” - suurepärane 91-100%
„B” - väga hea 81- 90%
„C” - hea 71- 80%
„D” - rahuldav 61-70%
„E” - kasin 51- 60%
„F” - puudulik 0 – 50%

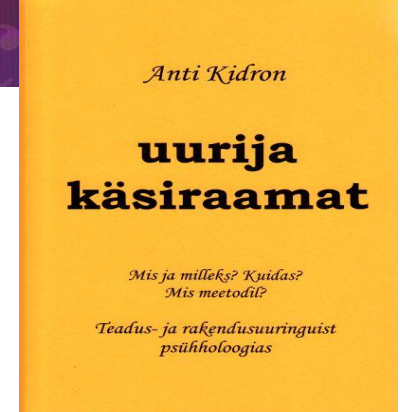
Eksamile saab kaasa võtta A4 paberi. Järeleksam on ilma lisamaterjalide kasutamisetä. Praktilist tööd tuleb kaitsta eksamipäeval (lepime ajad kokku).

Lisalugemist

- “Uuri ja kirjuta”
 - Autorid: S. Hirsjärvi, P. Remes, P. Sajavaara
 - Kirjastus: Medicina
 - Ilmumisaasta: 2005



- “Uurija käsiraamat”
 - Autor: A.Kidron
 - Kirjastus: Mondo



Lisalugemist

- “Andmete analüüs ja tõlgendamine sotsiaalteadustes”
 - Autor: Liina Mai Tooding
 - Kirjastus: Tartu Ülikooli kirjastus
 - Ilmumisaasta 2015
- “Andmeanalüüs sotsiaalteadustes”
 - Autor: Liina Mai Tooding
 - Kirjastus: Tartu Ülikooli kirjastus
 - Ilmumisaasta: 1998



Kvalitatiivne vs kvantitatiivne

Mis/kes?

Mil määral?

Kuidas?

Mis ulatuses?

Missugune?

Kui sageli?

Aga on ka küsimusi, mis sobivad hästi mõlemale poole
Praktikas vajalik **kombineerimine**: kahe meetodi triangulatsioon,
kvalitatiivselt leitu kinnitamine kvantitatiivse uuringuga või
kvantitatiivselt leitu täpsustamine kvalitatiivuuringuga

Empiiriline uuring

- Järjesta empiirilise uuringu etapid <http://lingid.ee/empiirilineuuring>

- Uurimisprobleem
 - küsimus, hüpotees, eesmärk
- Valikumeetod
 - juhuslik valim, üks juhtum, mitu juhtumit
- Andmekogumis meetod(id)
 - struktureeritud ankeet, struktureerimata intervjuu, ...
- **Andmeanalüüsi meetodid**
 - andmete korrastamine, statistilised meetodid, kodeerimine
- **Tulemused/järeldused**
 - kirjeldused, empiirilised üldistused, seaduspärasused,...

Valimi moodustamine

Tõenäosuslikud valimid
(*probability sampling*)

VEA HINDAMINE

Vea hindamine TÄPNE
Matemaatilised valemid
Tõenäosusteooria

**Viga hinnatakse suuremaks väiksemate
gruppide korral**

Vahemikhinnang, olulisustestid,
dispersioonanalüüs, regressioonanalüüs

Mittetõenäosuslikud valimid
(*non-probability sampling*)

VEA HINDAMINE

Viga hinnatakse suuremaks
kuna tüüpilised või mugavad objektid
võivad osutada ebatüüpilisteks või
mitteesinduslikeks.

Üks-ühele tehtavaid üldistusi
üldkogumile teha ei
saa. Võimaldab näha töö tulemustes
üldisemalt kehtivaid tendentse.

Andmete kogumine

Töös püstitatud hüpoteesi(de) kontrollimiseks on vaja andmeid, fakte.

otsene loendamine ja mõõtmine dokumenteerimine / koondandmete talletamine struktureerimata e avatud, poolstruktureeritud ja struktureeritud **intervjuud** (suulised küsitlused) struktureeritud ja poolstruktureeritud ankeedid (kirjalikud **küsimustikud**) (standardiseeritud) **võimekus-** ja **sooritustestid** psühholoogilisi ja sotsiaalseid nähtusi (nt hoiakud, väärtused, hinnangud, jms) mõõtvad **testid**

Küsimuste/mõõdikute kavandamisel mõtle ja otsi infot järgneva kohta:

- kas saab uuritavat nähtust mõõta **otse** või läbi indikaatori(te)?
- kas antud nähtust on **eelnevates uuringutes** mõõdetud ning kas vastav kirjandus on usaldusväärne? Kas saab juba **olemasolevaid mõõdikuid** kohandada või tuleb välja töötada uued mõõdikud/küsimused?
- kas peaks kasutama **piloteerimist** ning järelkontrolli?

Tunnuse tüübid

Arvestame **järjestatavust, skaalavahemike võrdsust**

Nimitunnused

Nimitunnuse väärtuseid **ei saa järjestada**

Järjestustunnused

Järjestustunnuste väärtuseid **saab järjestada**
Skaalavahemikud ei ole võrdsed

Intervalltunnused

Skaalavahemikud on võrdsed

Vähe erinevaid väärtuseid

Palju võimalikke väärtuseid

Binaarsed tunnused

Kaks võimalikku väärtust

Andmestiku korrastamine

- ülevaade puuduvatest väärtustest, otsus
- sisestusvigade otsimine ja parandamine
- skaalade pööramine
- kategooriate liitmine
- andmete grupeerimine
- andmesiltide korrastamine, tunnuse nimede korrastamine

Statistika jaguneb...

KIRJELDAV STATISTIKA

- Tabelid, diagrammid, seosed, kirjeldavad arvnäitajad, ...
- kirjeldav, võrdlev

ÜLDISTAV STATISTIKA

- Olulisustestid, vahemikhinnang, ...

TUNNUSTE/VÄÄRTUSTE KOONDAMINE

- Klasteranalüüs, faktoranalüüs, skaleerimine, diskriminantanalüüs, ...

Meetod sõltub analüüsiküsimusest ja andmetest

- tekst
- tabel
- diagramm

VÕRDLEV
(erinevused gruppide vahel)

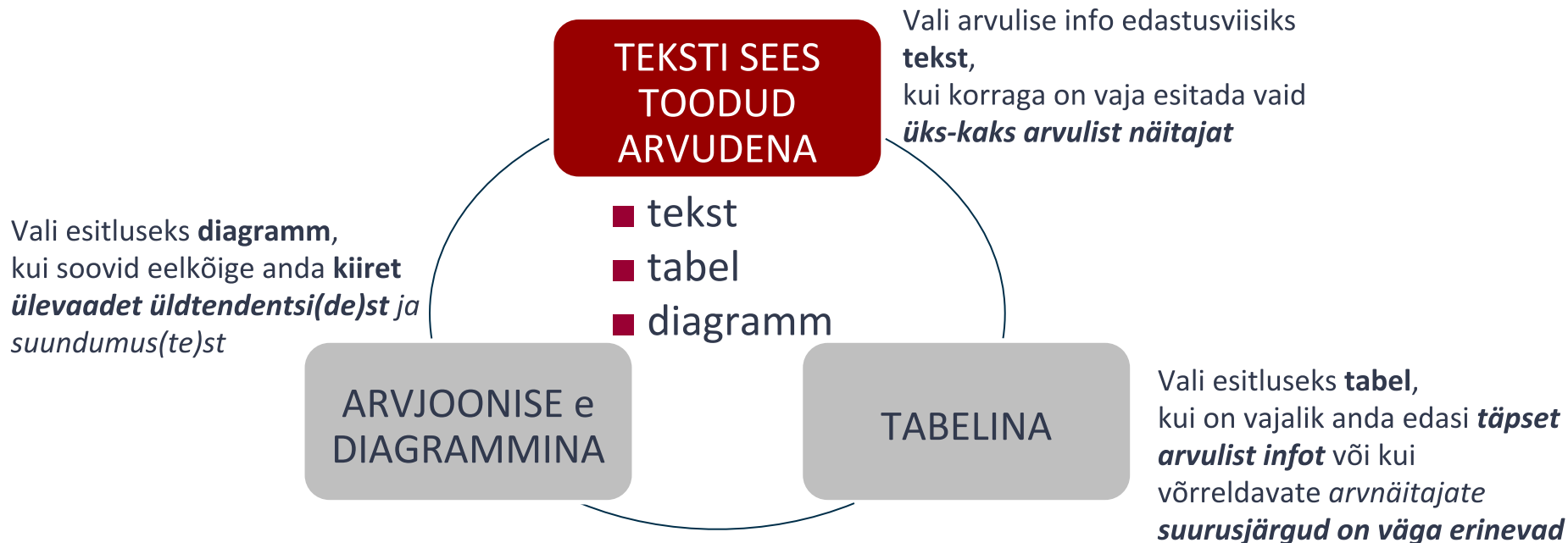
KIRJELDAV
(esmane analüüs)

- tekst
- tabel
- diagramm

KORRELATSIOON
(seosed tunnuste vahel)

- tekst
- tabel
- diagramm

Statistiliste andmete esitamine



Esitlusviis peaks toetama parimal viisil tulemuste sisust kiiret ja õiget arusaamist ning olema kompaktne.