

Kursus: Mitmemõõtmeline statistika

Seminar I: Mitmefaktoriline dispersioonanalüüs.

Õppejõud: Katrin Niglas
PhD, dotsent
informaatika instituut



Mitmefaktoriline dispersioonanalüüs (2-Way ANOVA, 3-Way ANOVA, ...)

- Ühe tunnuse keskmiste võrdlemine mitme grupeeriva tunnuse lõikes
- Üks sõltuv tunnus (arvtunnus)
- Mitu sõltumatut tunnust ehk **faktorit/faktortunnust**
(objekte gruppidesse jagavad tunnused ehk tunnused, millel on 2 või enam selgesti eristuvat **taset** st väärtust)
- **Kahefaktoriline** dispersioonanalüüs, **kolmefaktoriline** disp.anal. jne
- Faktori **peamõju**
- Faktore **koosmõju**
- **Täismudel** (sisaldab kõiki võimalikke mõjukomponente)



Mitmefaktoriline dispersioonanalüüs – hüpoteeside püstitus

- Testitakse rohkem kui ühte hüpoteeside paari (iga peamõju ja iga koosmõju kohta eraldi)!
- **Nullhüpoteesid:**
Erinevused üldkogumis /vastavate/ gruppide keskmiste tasemetel vahel puuduvad
- **Peamõju kirjeldavate testide sisukas hüpotees:**
Vähemalt ühel /vastava/ faktori tasemel e ühel /antud/ tunnuse poolt moodustunud grupil on üldkogumis teistest erinev keskmine tase ehk antud faktor e tunnus mõjutab tulemust (on tulemusega seotud)



Mitmefaktoriline dispersioonanalüüs – hüpoteeside püstitus

- **Koosmõju kirjeldavate testide sisukas hüpotees:**
Vähemalt ühe /antud/ faktorite tasemekombinatsioonide paari korral on keskmiste erinevus gruppide vahel teistest paaridevahelistest keskmiste erinevustest üldkogumis erinev ehk ühe faktori efekt/mõju tulemusele sõltub teise faktori tasemetest ehk ühe faktori efekt/mõju tulemusele on teise faktori poolt moodustatud gruppides erinev

NB! Faktor = sõltumatu tunnus

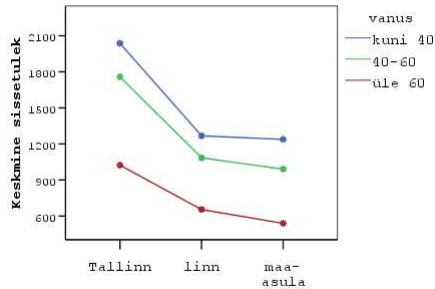
faktori tase = sõltumatu tunnuse väärtus = üks võrreldav grupp

PS! Koosmõju on statistilises mõttes sümmeetriline!

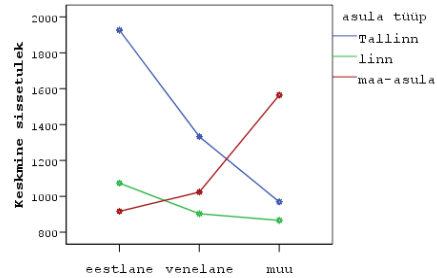


Koosmõju

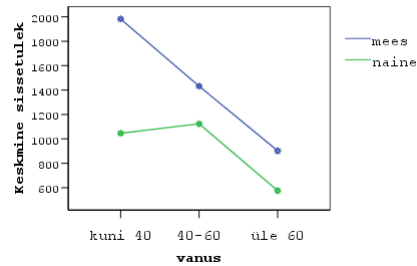
Vanusel ja asula tüübil koosmõju puudub



Asula tüübil ja rahvusel on koosmõju



Vanusel ja sool on koosmõju



Dispersioonanalüüs – eeldused

- **Sõltuv tunnus:** arvtunnus (st praktikas võrdsete vahedega skaala)
- **Sõltumatu(d) tunnus(ed):** väheste väärtustega nn grupeeriv(ad) tunnus(ed)
(NB! Sõltumatute ja sõltuvate gruppide jaoks eraldi testid!)
- **Normaaljaotus**
Mittetäidetuse korral mitteparameetriline test (nt Friedman või Kruskal-Wallis).
PS! Kui valimite suurus üle 20 siis tavaliselt selle eelduse suhtes robustne
- **Gruppide hajuvuste võrdsus (Levene's test)**
Mittetäidetuse korral mitteparameetriline test või parandatud F-statistik (Brown & Forsythe või Welch F)
PS! Mõnede allikate kohaselt robustne kuni kolmekordsete hajuvuse erinevuste suhtes.

Statistilise olulisustesti põhisammud:

E I: Analüüsisin olemasolevaid andmeid kirjeldava statistika meetodite abil ning leidsin midagi „huvitavat“ (nt. erinevuse või seose jne)

E II: Tekkis küsimus: „Kas võib üldistada?“

$\bar{U} \rightarrow V \rightarrow \bar{U}$ (v eksperimentaalne disain)

I. Õige olulisustesti valik (lähtuvalt probleemist ja andmetüübist)

II. Valitud olulisustesti eelduste kontroll:

ei
jah

III. Hüpoteesid: sisukas hüpotees H_1 :

nullhüpotees H_0 :

Olulisuse nivoo α

(„Kui väike peab olema H_0 kehtimise tõenäosus, et me võiks ilma suurema riskita ta mittekehtivaks tunnistada?“)

IV. Arvutused

eesmärgiks hinnata H_0 kehtimise tõenäosust p

(„Kui suur on tõenäosus, et olukorras, kus H_0 kehtib, tekkis valmis olnud erinevus v seos juhuse tõttu?“)

($p = Sig = olulisuse\ tõenäosus$)

V. Otsus tulemuse kohta:

$p > \alpha$ H_0 jääb kehtima - statistiliselt mitte oluline (ei üldista)

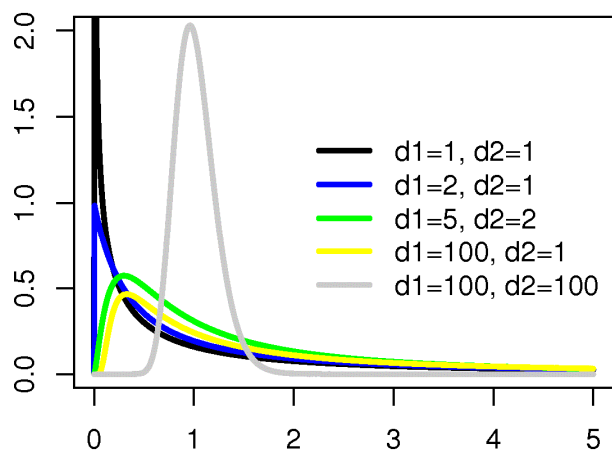
$p \leq \alpha$ H_1 tõestatud - statistiliselt oluline (võib üldistada)



TALLINNA ÜLIKOOL

VI. Järelduse sõnastamine

F-jaotuste pere



TALLINNA ÜLIKOOL