I PRAKTIKUM: SISSEJUHATUS PRAKTILISSE ANDMETÖÖTLUSESSE

Esimene praktikum on sissejuhatava loomuga – käsitleme põgusalt erinevaid võimalusi, mis vahenditega andmeid töödelda, tutvume statistikapaketiga JASP ning viime läbi lihtsamad andmeanalüüsid.

Käesoleva praktikumi eesmärgiks on (a) anda lühike ülevaade saadavalolevatest andmeanalüüsipakettidest ning –tarkvarast, (b) tutvuda JASP-i keskkonnaga ning (c) läbi viia mõningad andmetöötluse operatsioonid.

1. STATISTIKATARKVARAST ÜLDISELT

Uurimisküsimusele vastamiseks ei saa ära jätta andmeanalüüsi osa – statistikatarkvara kasutatakse keerulisemate arvutuste tegemiseks ning mudelite loomiseks. Ehkki kõiki neid arvutusi on sisuliselt võimalik läbi viia nn „käsitsi“ (arvutamine paberi-pliiatsi, kalkulaatori abil), ei ole see tänapäeval enam otstarbekas, kuivõrd arvutite ning andmetöötlustarkvara kasutamine võimaldab tohutut kokkuhoidu ressursside, nagu aeg ning raha, arvelt.

Inimesed võivad igapäevaelus üsna lihtsasti leida keerulisemate andmeanalüüside vilju – folkloori kohaselt enim tarvitet jututeema – ilm – on üks nendest näidetest. Ehkki inimesed on püüdnud tuhandeid aastaid ilma ennustada, ei ole see olnud kunagi varem nii hõlpsasti ja kõrge täpsusega teostatav, kui on tänapäevased, arvutite abil tehtud mudeldused.

Niisamuti on inimkäitumise, mõtlemise ning tundmise uurimisega – kui umbes sajand tagasi viidi läbi tunde, päevi ja nädalaid kestnud arvutusi, siis tänapäeval on võimalik kõike seda tihtipeale teha lausa minutite ja mõne hiireklikiga!

Iseküsimus on aga see, millist tarkvara on kasutatud ennustusmudelite loomiseks. Ehkki tasub meeles pidada, et **lõppkokkuvõttes on (kvaliteetne) andmestik iseenesest kõige tähtsam** osa andmeanalüüsist (ja ka uuringust endast), ning et lihtsamaid arvutusi saab teha sisuliselt kõigi statistikaprogrammidega[1], on tänapäevased statistikapaketid tihtipeale modifitseeritud kasutajakesksemaks, st ilmaennustuste jaoks on meteoroloogidele mugavdatud programmid (nt WXSIM, jt) ja ka nt sotsiaalteadlastele on mugavaid vahendeid, mis ei pruugi eeldada matemaatilist tausta (nt JASP, SPSS, Stata, jt).

Samas tasub aga teada, et suur osa kasutatavast statistikatarkvarast on tihtipeale tasuline[2]. Ent on ka suur hulk tasuta, vabavaralisi pakette – reeglina ei pruugi nendes saada läbi viia keerukamaid analüüse, mida võimaldavad tasulised ning enamkasutatavad programmid. Üheks näiteks on SPSS-i vabavaraline analoog PSPP.

Sotsiaalteaduste-inimestele on tuttavamateks nimedeks Microsoft Excel, Stata, programmeerimiskeel R, SPSS ning JASP. Milline neist on parim? Suuresti on tegemist isikliku eelistuse küsimusega – igal variandil on omad eelised, aga ka omad kitsaskohad. Sotsiaalteadustes kasutatakse enamasti SPSS-i, kuid üha rohkem populaarsust kogub vabavaraline programm JASP[3]. Samuti kasutatakse programmeerimise põhimõtetele toetuvat R-i. Käesoleva ainekursuse põhiliseks tööriistaks on JASP. Selle tarkvara kasuks räägivad mitu faktorit: 1) tegemist on vabavaralise programmiga; 2) programmi kasutamine on tehtud võimalikult lihtsaks; 3) programm võimaldab teha kõiki peamisi sotsiaalteadustes vajaminevaid andmeanalüüse. Lisaks on tegemist uue ja kiirelt areneva programmiga, mistõttu on internetiavarustest hõlpsasti leitav palju abimaterjale ja juhendeid programmi erinevate võimaluste kasutamiseks.

2. TUTVUMINE MS EXCELI VÕIMALUSTEGA

JASP on programm, mille abil teostada andmeanalüüse, kuid eelnevalt tuleb andmed sisestada andmetabelisse. Selleks on sobilik näiteks Excel või mõni muu analoogne tabeltöötluseks mõeldud programm.

MS Excel on küllalt laialdaselt kasutatud programm, millega saab muuhulgas ka andmeanalüüsi teostada. Ehkki keerulisemate mudeldamiste läbiviimine võib Excel-is olla tülikas, saab selles täiesti edukalt lihtsamaid statistilisi operatsioone teostada. Mis kõige tähtsam, Excel on mugav keskkond, kus oma andmetabelid enne statistiliste

[1] Erinevate statistikapakettide võrdlus: mida saab ning mida ei saa teha: https://en.wikipedia.org/wiki/Comparison_of_statistical_packages

[2] 50 vabavaralise statistikatarkvara nimistu: http://www.predictiveanalyticstoday.com/top-free-statistical-software/

[3] JASP-i ja SPSS-i võrdlus: https://jasp-stats.org/2017/11/01/jasp-vs-spss/

analüüside läbiviimist kerge vaevaga sobivaks toimetada. Selles praktikumis vaatamegi esialgu põgusalt üle ka Exceli võimalused.

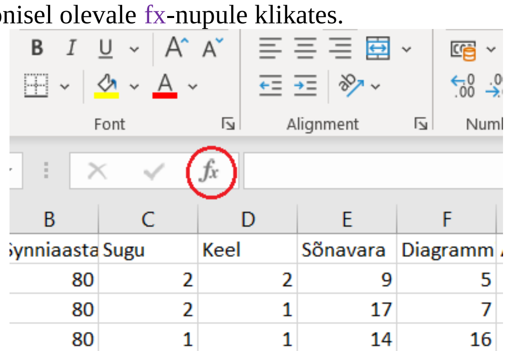
Lisamärkus

Kellel pole töötavat Excelit oma arvutis olemas, siis tudengitel on võimalik tasuta versioon alla laadida selle juhendi abil: https://wiki.ut.ee/display/AA/Microsoft+Office+365+Educationi+paigaldamine Selle kasutamine pole siiski kohustuslik ning sama hästi peaks töötama ka teised tabeltöötlusprogrammid (näiteks .ods vorminguga).

Käesoleva praktikumi põhifailiks on **test11.xls**. Salvestame faili käsklusega *Save As...* ning paneme faili nimeks **test11Ex.xls**. Harjutused Exceliga võib salvestada selle nimega faili.

Vaatame, kuidas viia läbi muutujate järjestamine. Selleks tuleb esmalt vastava muutuja andmepunktide tulp teha „aktiivseks". Menüüribal asuva *Home* all on paremapoolses osas käsklus *Sort & Filter*, millele klikates saate valida *Sort Smallest to Largest*. Sellele käsklusele klikates ilmub tõenäoliselt aken, mis küsib, kas „aktiivseks" valitud andmetulpa tuleb kohelda muust andmestikust iseseisvana või mitte. Valige *Expand the Selection*, mis tähendab seda, et ka kõik teised muutujad on indiviididega seotult võetud ning vastavalt järjestatud.

Et leida kõik statistikud, tasub iga statistiku jaoks kasutada eraldi ruutu (vt andmestiku alumises osas). Excelis on funktsioonid, mis on leitavad joonisel olevale fx-nupule klikates.



Avaneb menüü, milles saate valida vajamineva funktsiooni. Kui olete funktsiooni ära valinud, küsib Excel, mis andmestikule arvutused rakendatakse. Tehke aktiivseks indiviidide vanuse kohta infot kandev tulp „Synniaasta" (meie andmestikus: B2-B1351); alternatiivselt võite ka avanenud lahtrisse kirjutada kasutatava andmestiku vahemiku (nt: B2:B1351). Kas statistik ilmus kastikesse?

On ka lihtsamaid viise, kuidas kõiki kirjeldavaid statistikuid *korraga* arvutada. Selleks tuleb järgida käsklusterida: *File – Options – Add-Ins –* valida kasti alumisest osast *Excel Add-ins* ning *Go*, misjärel avaneb aken nimega *Add-Ins*. Sealt valida *Analysis ToolPak* ning klikata *Ok*.

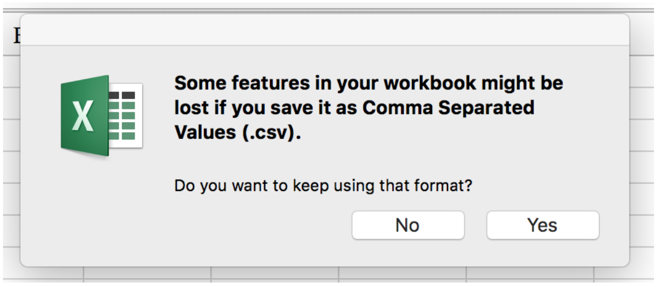
Menüüriba käskluse *Data* alla paremale ilmus *Data Analysis*. Kui klikite selle peale, valige nimistust *Descriptive Statistics*. Excel soovib, et sisestaksite andmestiku ulatuse, mille peal arvutused läbi viiakse. Valige muutuja *Synniaasta* andmestik. Samuti valige *Summary Statistics* – et saaksite meid huvitavad kirjeldavad statistikud. Lisaks saate täpsustada, kuhu kuvatakse tulemuste tabel. Vaikimisi tekitatakse tavaliselt uus Excel-i tööleht.

Lisamärkus

Kui teil on enda kogutud andmed, siis tasub andmed sisestada Excelis ja seejärel salvestada failist üks versioon csv-formaadis, mida hakata kasutama JASP-is analüüside tegemiseks. JASP-is saab samuti andmevälju muuta, kui teete topeltklõpsu andmeväljale. Seejärel toob JASP esile andmefaili Excelis, kus saate teha vajalikke muudatusi. Näiteks kustutada andmetabelist üleliigsed veerud või vajadusel muuta mõne lahtri väärtust. Pärast andmefaili salvestamist kanduvad muudatused üle JASP-is olevasse andmefaili.

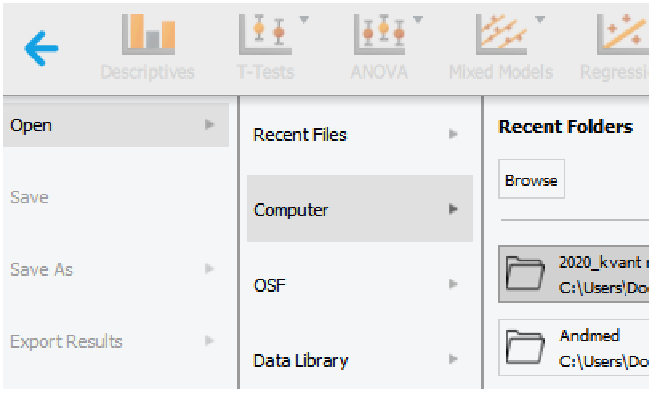
3. SISSEJUHATUS JASPI

Võtame uuesti algfaili **test11.xls** ning salvestame selle nimega **test11JASP.csv**. Pange tähele, et see fail peaks olema nüüd teistsuguses formaadis, vastasel juhul me ei saa faili JASP-iga avada. Salvestame faili käsklusega *Save As...* ja valime formaadiks **CSV (komaga eraldatud).**
Tõenäoliselt näete ka järgnevat akent.



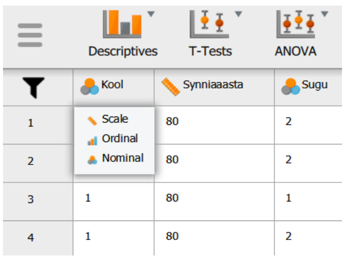
Vajutage „jah“.

Järgnevalt tahame avada seda faili JASP-is. Selleks avage programm ja valige *File* - *Computer* - *Browse* ning otsige eelnevalt salvestatud fail üles.



Muutujate kontrollimine

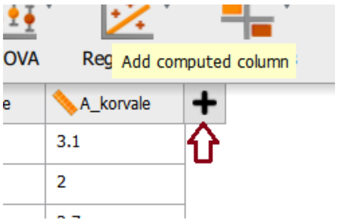
Esmalt tuleks kontrollida, kas andmefailis olevad muutujad on JASP-is õiget tüüpi. Selleks kontrollige, et iga muutuja nime ees olev sümbol vastaks muutuja skaalale. Tüübi muutmiseks tuleb vajutada sümbolile.



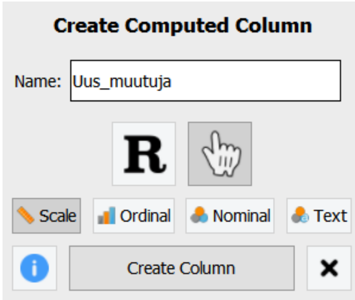
Uue muutuja loomine

JASP-i sees muutuja loomine

JASP-i sees uue muutuja arvutamiseks vajutage muutujate nimede kõrvale olevale pluss märgile:



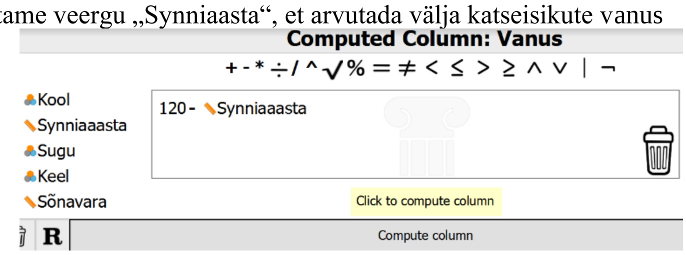
Järgnevalt tuleb ekraanile aken, kus tuleb teha järgmised sammud: 1) sisestada soovitud muutuja nimi, 2) teha valik, kas arvutate uue muutuja R süntaksi abil või lohistades (meie kasutame praktikumides teist varianti), 3) mis tüüpi muutujaga on tegemist (pidev, ordinaalne, nominaalne või tekst). Kui need valikud tehtud, vajutage all olevat nuppu „Create column".



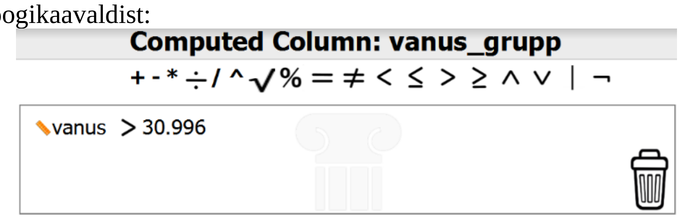
Seejärel ilmub andmestiku üleval kast nimega „Computed Column: „sisestatud nimi". Sinna kasti saate lisada uue veeru arvutamiseks vajaliku tehte. Uute veergude arvutamiseks saate kasutada tavalisi arvutustehteid (liitmine, lahutamine, jagamine, korrutamine) või loogikaavaldisi (mingi teise tulba väärtus on suurem kui meie etteantud väärtus), mis on välja toodud kasti kohal. Kastist vasakul näete oma muutujate nimekirja ning kastist paremal on erinevad tehted, mida on saab lisaks võimalik kasutada (nt summa, keskmine, logaritm jpm).

NB JASP ei luba matemaatilisi tehteid teha diskreetse muutujatega (st nominaalsed ja järjestusskaalal)!

1) Lahutamine: kasutame veergu „Synniaasta", et arvutada välja katseisikute vanus



2) Loogikaavaldis: teeme kaks gruppi vanuse alusel. Selleks arvutame vanuse keskmise väärtuse ja kasutame tehtena loogikaavaldist:



See tehe loob uue muutuja, kus kõik katseisikud vanemad kui 30.996 aastat on grupis 1 (TRUE) ja kõik nooremad kui 30.996 on grupis 0 (FALSE).

Muutujate või tehete eemaldamiseks tuleb need lohistada prügikasti.

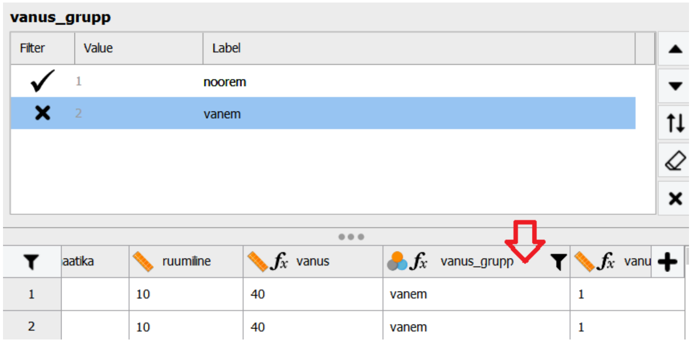
JASP-ist väljas muutuja loomine

JASP-i sisese funktsionaalsuse kasutamine võib olla natuke piiratud, eriti juhul kui tahate jaotada pideva tunnuse kolmeks või rohkemaks kategooriaks (nt. vanuse alusel kolme kategooria tegemine: noored, keskmised, vanad). Lisaks JASP-is olevale funktsionaalsusele saate kasutada arvutis olevat andmetöötlusprogrammi (tavaliselt Excel). Vajutage JASP-is kaks klikki mõnele väljale andmestikus. Seejärel avaneb andmetöötlusprogramm koos JASP-is olevate andmetega. Kui teete seal muudatusi ja seejärel need salvestate, siis kanduvad need muudatused koheselt JASP-is olevale andmestikule üle.

Vaadake selle kohta lähemalt https://www.youtube.com/watch?v=1dT-iAU9Zuc.

Andmete välja filtreerimine

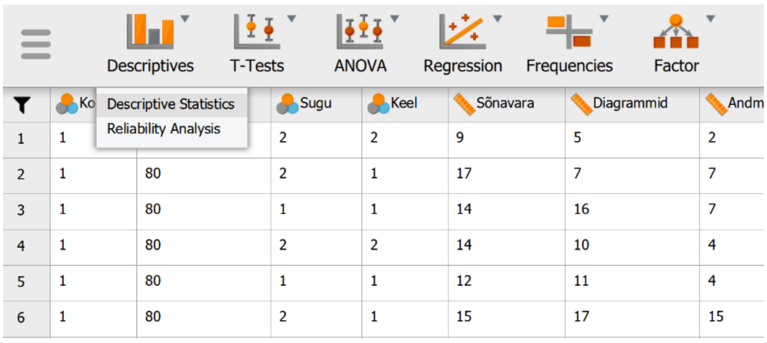
Vahel võib olla tarvis osad andmed mõnest analüüsist välja jätta. Näiteks tahame teha analüüse ainult noorema vanusegrupiga. Selleks on mitu moodust. Ühe võimalusena on muidugi vanema vanusegrupi andmed tabelist kustutada – seda tavaliselt siiski ei tasu teha, kuna neid võib hiljem veel vaja minna. Parem mõte on vanema vanusegrupi andmed ajutiselt andmetest välja filtreerida. Selleks tuleb vajutada muutuja nimele, mille peale avaneb filtreerimisaken – tunnus, mida soovite eemaldada, tuleb ristiga ära märkida. Pärast edukalt filtreerimist näete välja filtreeritud andmeid andmetabelis hallina.



Keerukamateks filtreerimisteks saab kasutada andmetabeli vasakul nurgas olevat filtreerimise nuppu ning sisestada loogikaavaldise kujul filtreerimisest pääsemise tingimus (näiteks „vanus < 30“ jätab analüüsidesse alles ainult need isikud, kes on nooremad kui 30).

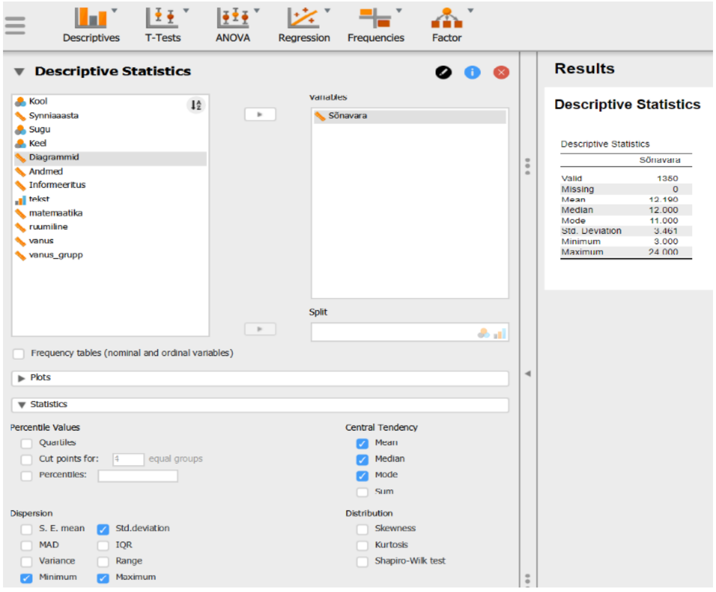
Kirjeldav statistika

Andmetega tutvumiseks on hea alustada mõningate kirjeldavate statistikute arvutamisest ja jooniste tegemisest. Selleks valige ülemisest menüüst *Descriptives - Descriptive Statistics*.



Järgmisest aknast valige muutuja(d), mida soovite analüüsida. Täpsemalt saate statistikuid valida menüüst „Statistics“, kus näete kõige olulisemaid kirjeldavaid statistikuid: keskmine (*mean*), mediaan (*median*), mood (*mode*), haare (*range*), dispersioon (*variance*), standardhälve (*standard deviation*), asümmeetriakordaja (*skew*) ja järsakusaste (*kurtosis*).

Valige nõutud statistikud ja viige läbi analüüs. Seejärel peaks avanema paremal analüüsiaken, kus näete arvutuste tulemusi:

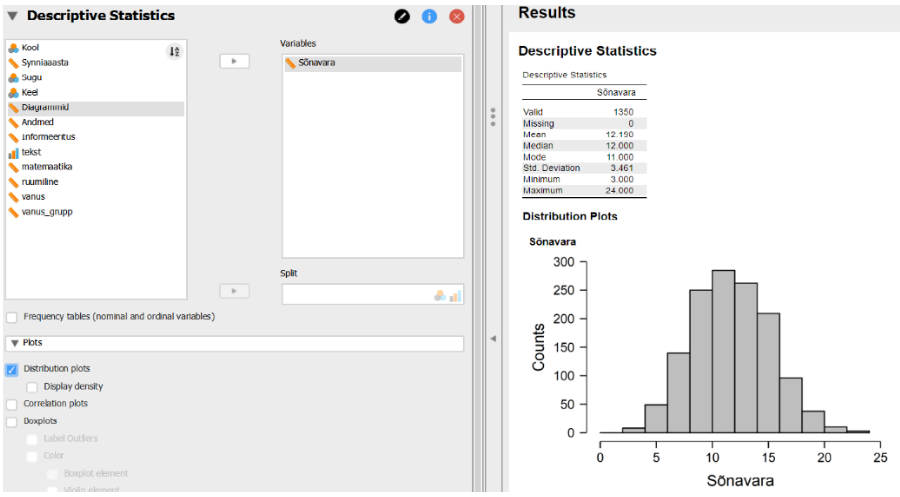


Lisamärkus

Kirjeldavat statistikat on võimalik vaadata ka mitme eraldi grupi lõikes. Selleks tuleb aknasse „Split" lisada mõni kategooriline tunnus, näiteks sugu.

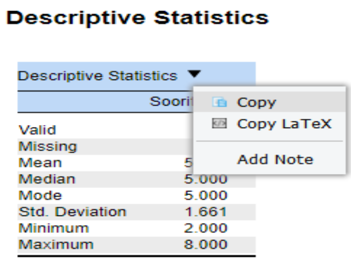
Histogrammi tegemine

Histogrammi tegemiseks avage uuesti kirjeldava statistika menüü: *Descriptives - Descriptive Statistics*. Valige alamenüüst *Plots Distribution Plots*. Seejärel ilmub automaatselt analüüsiaknasse valitud muutuja histogramm.



Lisamärkus:

> *JASP võimaldab kopeerida tabeleid ja jooniseid Wordi faili. Kopeeritud tabelit saate wordis vastavalt vajadusele (APA stiilile) kohandada.*



Sarnaselt tavalise tabeliga te saate loodud histogrammi salvestada enda Wordi faili. Samuti saate salvestada JASP-is tehtud töö .jasp formaadis, mis säilitab tehtud analüüsid. Selleks valige: File - Save As - Computer – Browse.

4. ÜLESANDED ISESEISVAKS HARJUTAMISEKS

Avage fail „P1_eksperiment.csv“ – selles failis on kasutatud andmeid Dmitri Rozgonjuki seminaritöö (2013)[4] andmestikust. Selles uuringus sooritasid katseisikud verbaalset ülesannet ning inimestele anti erinevat tagasisidet ja mõõdeti selle mõju enesehinnangu näitajatele Andmefailis on 7 muutujat: katseisiku kood, vanus, eriala, katsegrupp (BFLPE), testi soorituse järel sooritusele antud hinnang (SooritusEH), soorituse järel hinnatud üldine verbaalne võimekus (ÜldineEH) ning kommentaarid testi/uuringu kohta.

Sooritage järgnevad ülesanded, kasutades selleks programmi JASP:

1) Leia uuringus osalenute keskmine vanus, vanuse mood ning mediaan. Kui vanad olid kõige noorem ja kõige vanem osaleja?
2) Joonista vanuse jaotustabel ehk histogramm
3) Pange katsetingimuse (tulp “Tagasiside”) väärtustele sildid: (1=positiivne, 2=negatiivne)
4) Tehke uus muutuja nimega “EHindeks”, mis kujutab endast kahe enesehinnangu keskmist (SooritusEH ja ÜldineEH)
5) Kustuta andmetabelist analüüsiks ebavajalikud read – millised ja miks?
6) Lisaküsimus: Milline veerg kujutab selles katses sõltumatut muutujat? Milline veerg on sõltuv muutuja?

[4] Seminaritöö on saadaval TÜ digitaalses andmebaasis DSpace: http://goo.gl/UtVuQr

II PRAKTIKUM: PÕHILISED KIRJALDAVAD STATISTIKUD JA NORMAALJAOTUSE TESTIMINE

Selles praktikumis kordame (1) kuidas andmefaile (.csv) avada ja (2) kuidas JASPis kirjeldavaid statistikuid tellida. Selles praktikumis vaatame lisaks, (3) kuidas luua tunnus, mis jaotab andmed mitmeks sarnase suurusega grupiks, (4) õpime tunnuste jaotuste normaaljaotusele vastavust hindama, (5) õpime tunnuseid standardiseerima ja (6) võrdlevaid karpdiagramme joonistama ning grupisiseseid hälbivaid väärtusi leidma ja tuvastama, mis juhatab sisse järgmise praktikumi keskmiste võrdlemise temaatika.

Praktikumi põhiandmestikuks on esimeses praktikumis kasutatud fail p1_AKtest.csv.

1. ANDMETE AVAMINE

Andmefaili avamiseks avage kõigepealt JASP ja seejärel valige menüüst (kolm sinist horisontaalset joont üleval vasakus nurgas): *Open > Computer > Browse*. *Browse* nupule klõpsates avaneb lehitsemisaken, milles navigeerige kausta, kuhu andmefaili salvestasite. Andmefail peaks olema csv formaadis. Kui teie andmefail on mõnes muus formaadis, siis avage andmefail kõigepealt Excelis ja salvestage andmed csv formaadis (*File > Save As > Save* (enne *Save* vajutamist valige selle ees olevasse lahtrisse ka õige failiformaat ehk CSV (comma delimited)).

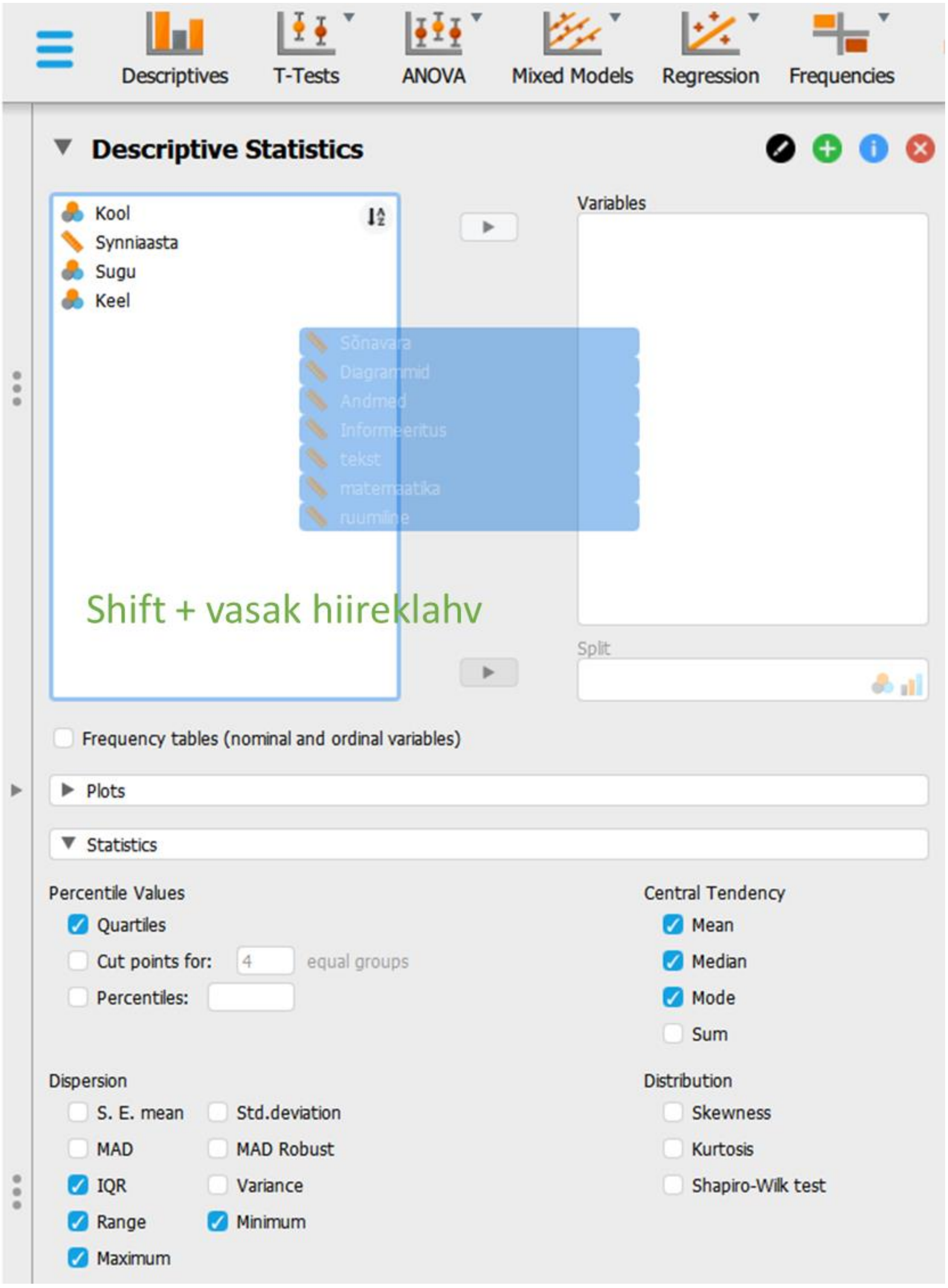
2. KIRJELDAV STATISTIKA

Ülesanne 1. Arvutage akadeemilise testi erinevate alaskaalade keskmine, mood, mediaan, maksimaalne, minimaalne väärtus, haare, kvartiilhaare ja kvartiilid. Tekitame iga alaskaala kohta sagedusjaotuse joonise.

Selleks klõpsake *Descriptives* ikoonil. Avanenud *Descriptive Statistics* aknas lohistage vasakul paiknevad alaskaalade nimetused paremal asuvasse kasti nimega *Variables* (mitme tunnuse korraga valimiseks tehke hiireklõps esimesel tunnusel (nt Sõnavara), mille soovite aktiivseks teha ja seejärel *Shift* klahvi all hoides vajutage viimasel tunnusel (nt ruumiline), mida sooviksite aktiivseks muuta). *Statistics* menüüs tehke linnuke soovitud kirjeldavate statistikute ette.

Kui andmed on intervall või suhte skaalal, siis pakub histogramm hea võimaluse kogutud andmete esmaseks inspekteerimiseks. Sagedusjooniste ehk histogrammide saamiseks klõpsake lahti *Statistics* menüü üleval olev *Plot* menüü. *Plot* menüü alaosas *Basic plots* tehke linnuke *Distribution plots* ette.

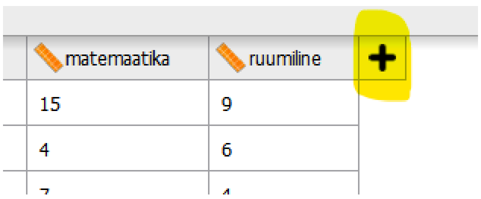
Kui teha linnuke ka Display density ette, siis ilmub joonisele tume joon, mis väljendab tunnuse hinnatud jaotumist pideva joonena. Tegemist on hinnanguga, sest iga joont moodustava punkti saamiseks võetakse paljude vaatluste kaalutud keskmine selliselt, et punkti lähedal paiknevad vaatlused panustavad saadud väärtusesse enam kui sellest eemal paiknevad vaatlused (vt ka Kernel smoother).



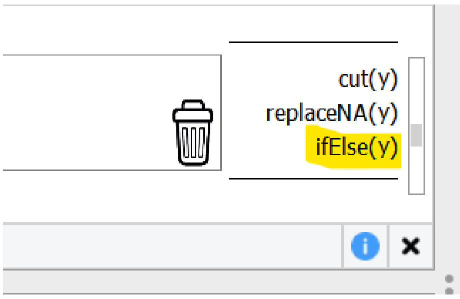
Ülesanne 2. Viiekümnes protsentiil on identne ühega järgnevatest: mood, mediaan, aritmeetiline keskmine?

Ülesanne 3. Tekitage andmetabelisse veerg, mis jaotab katseisikud sõnavara testi alusel kolmeks grupiks.

Selle ülesande lahendamiseks looge uus muutuja. Uue muutuja loomiseks vajutage muutujate kõrval oleval pluss märgil.



Avanenud aknas andke muutujale nimetus, muutke muutuja skaala tüüp mittepidevaks ja vajutage *Create Column* nupule). Parempoolsest ja prügikasti kõrval paiknevast keritavast menüüst otsige üles funktsioon ifElse(y) ja tehke sellel klõps.



Seejärel ilmub muutujate kõrval paiknevasse kastikesse tekst ifElse(*test*, *then*, *else*), kus test tuleks asendada loogikaavaldisega ning *then* ja *else* ühe või teise tagajärjega (näiteks sildiga). Järgnevalt tuleb pisut põhjalikumalt juttu „if else“ tüüpi süntaksi loogikast. Näiteks kui tahaksime Sõnavara testi tulemuste alusel tekitada kaks gruppi, siis võiksime kirjutada järgmiselt ifElse(Sõnavara < 15, 1, 2). Sellisel juhul annaks programm kõigile alla 15 punkti jäävatele vaatlustele sildi „1“ ja kõigile teistele vaatlustele (mis seda võrdsust ei rahulda) sildi „2“. See on selle pärast, et asendasime ifElse() funktsioonis nimetused *then* ja *else* numbritega „1“ ja „2“. Numbrite asemel võiksime kasutada ka verbaalseid silte: näiteks „alumine“ ja „ülemine“ (st ifElse(Sõnavara < 15, alumine, ülemine)) ja sellisel juhul annaks programm kõigile alla 15 punkti jäävatele vaatlustele sildi „alumine“ ja kõigile teistele „ülemine“ ehk teiste sõnadega funktsiooni loogika meie poolt antud nimetustest ei muutu. Tegelikult võiksime need sildid asendada ka mõne funktsiooniga (näiteks juhul kui sooviksime enne siltide andmetabelisse salvestamist veel mõne toimingu teha). Selleks, et tekitada uus muutuja, mis andmestiku mõne teise muutuja (nt keeletesti tulemuse) alusel kolmeks jagaks peaksimegi kasutama justnimelt seda võimalust. Õnneks saame jätkata juba olemasoleva süntaksiga (ifElse(Sõnavara < 15, 1, 2)), kus hetkel saavad kõik vaatlused, mis jäävad alla 15 punkti külge sildi „1“ ja ülejäänud sildi „2“.

Selles ülesandes tahaksime teise sildi alla jäävad vaatlused jagada veel kaheks osaks. Selleks asendame meie poolt eelnevalt loodud süntaksis arvu 2 ifElse() funktsiooniga ehk saame ifElse(Sõnavara < 15, 1, ifElse(test, then, else)). Teame, et vaatlused, mis number kahe asemel paiknevasse ifElse() funktsiooni jõuavad ei saa olla väiksemad kui 15, sest need sõeluti eelmise loogikaavaldise (Sõnavara < 15) poolt välja. Seega saame tekitada uue ja eelmisega analoogse loogikaavaldise, mis jaotab ülejäänud andmed kaheks osaks nt ifElse(Sõnavara < 16, 2, 3) ehk kogu süntaks tervikuna näeks välja järgnev ifElse(Sõnavara < 15, 1, ifElse(Sõnavara < 16, 2, 3)). Selliselt toimides saamegi tunnuse, mis jagab katsealused kolme gruppi:

a) need, kelle sõnavara testi tulemus oli väiksem kui 15;
b) need, kelle sõnavara testi tulemus oli suurem kui 14 ja väiksem kui 16 ja
c) need, kelle testitulemus oli 16 või suurem.

Selles näites valisime gruppideks jaotamise kriteeriumid juhuslikult, kuid päriselulise stsenaariumi korral peaksid gruppideks jaotamise kriteeriumid paremini läbimõeldud olema. Näiteks on praeguse kriteeriumiga saadud grupid väga ebaühtlase suurusega, mis võib nende võrdlemise väga keeruliseks teha. Ühtlase suurusega gruppide saamiseks saab edukalt kasutada protsentiile. Eelnevalt meenutasime, et mediaan vastab viiekümnendale protsentiilile, mis tähendab, et sellest allapoole ja ülespoole jääb võrdne hulk vaatlusi. Seega kui tahaksime andmefaili kaheks osaks jagada oleks mediaan heaks äralõikepunktiks. Kolmeks sarnase suurusega grupiks jaotamisel võiksime äralõikepunktidena kasutada aga protsentiile 33.3 ja 66.7, sest neist väiksemaid väärtusi on vastavalt umbes 33 ja 67 protsenti, mis jätaks kolmandasse gruppi umbes 33 protsenti mõõtmistest. Õnneks teeb JASP nende väärtuste leidmise väga lihtsaks. Selleks avame *Descriptives* menüü ja seal avame omakorda *Statistics* menüü. Esimese alajaotuse ehk *Percentile Values* all on valik *Cut points for:*, mille järele saab sisestada soovitud gruppide arvu ja programm arvutab välja sobivad protsentiilid. Võiksite proovida muuta eelmist süntaksi selliselt, et asendate selles juhuslikult valitud äralõikepunktid protsentiilidega. Selleks lohistage *Variable* aknasse muutuja *Sõnavara*. Seejärel arvutage muutujale protsentiilid selliselt, et need moodustaksid kolm sarnase suurusega gruppi (*Cut points for:* järele tuleks kirjutada arv kolm). Nüüd võtke saadud kaks väärtust ja asendage eelmises süntaksis 15 ja 16 nende väärtustega. Pärast seda tõsta uus muutuja *Variables* kasti all paiknevasse *Split* aknasse ja võrdle gruppide suuruseid. Ilmselt märkad, et grupid on küll sarnase, kuid mitte identse suurusega. See tuleneb sellest, et täpselt sama tulemuse võis saada terve hulk testitäitjaid ja nende erinevate protsentiilide alla liigitamine on mõnevõrra kokkuleppeline. Näeme näiteks, et 33.3 protsentiilist 44. protsentiilini on kõigi vaatluste väärtuseks 11. Seega pole antud muutuja puhul vahet, kas küsida 34. või 43. protsentiili, sest arvutatud väärtuseks tuleb ikka 11.

3. STATISTILISED MOMENDID JA NORMAALJAOTUS

Mõningaid jaotuse kirjeldamiseks kasutatud kirjeldavaid statistikuid nimetatakse ka momentideks. Eelnevalt leidsime aritmeetilise keskmise. Aritmeetilist keskmist kutsutakse ka esimest järku momendiks. Teiseks momendiks on hajuvus, kolmandaks asümmeetria ja neljandaks järsakus. Need nimetused pole päris juhuslikud, sest kõik neli statistikut on omavahel tihedasti seotud. Üldistatult võib öelda, et kõik nad on variatsioonid järgmisest valemi kujust:

$$m_k = n^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^k ,$$

kus k tähistab momendi järku, n vaatluste arvu, x_i i-ndat vaatlust ja $\bar{x}$ aritmeetilist keskmist. Selle üldistatud kuju mainimise idee on demonstreerida asjaolu, mis esmapilgul keeruliselt kõlavate terminite tagant kätte ei pruugi paista ehk seda, et need kirjeldavad statistikud kuuluvad samasse perekonda. Sellele vaatamata saab nad algvalemisse pisikeste variatsioonide tekitamisega panna väljendama väga erinevaid nüansse, kuid nende väärtuste leidmise aluseks olev matemaatiline loogika on väga sarnane. Kui esimest ja teist järku momendid (keskmine ja hajuvus) aitavad hinnata muutuja tüüpilist väärtust ja seda kui hästi see tüüpiline väärtus kõiki mõõdetud juhtumeid iseloomustab ehk hajuvust keskmise ümber, siis kolmandat ja neljandat järku momendid on abiks andmete normaaljaotuslikkuse hindamisel.

Ülesanne 4. Leidke kõigi alatestide tulemuste jaoks järgmised suurused: aritmeetiline keskmine, standardhälve, kvartiilid, järsakus ja asümmeetria. Samuti joonistage histogrammid ja vastake järgmistele küsimustele

a) Millise alatesti asümmeetria on kõige suurem?
b) Millise alatesti järsakus on kõige suurem?
c) Milline alatestistest vastab kõige paremini normaalkõverale?

Et välja selgitada, kas andmed pärinevad normaaljaotusega populatsioonist, saab läbi viia normaaljaotust uuriva testi. Selleks tehke linnuke *Shapiro-Wilk test* ette (*Descriptives > Statistics > Distribution > Shapiro-Wilk test*). Teine võimalus on kasutada hinnangu alusena asümmeetria- ja järsakuskordajat.

Pange tähele, et kui Shapiro-Wilki testi olulisuse tõenäosus on ***väiksem kui 0.05, siis ei ole testi kohaselt andmed normaaljaotuslikud****! Selle testi p-väärtuse tõlgendamine on levinud segaduste allikas. Testi olulisuse tõlgendamise loogika seisneb selles, et vaikimisi eeldame, et muutuja jaotus ei erine oluliselt normaaljaotusest. Test hindab, kas meil on piisavalt tõendeid, et see väide ümber lükata. Kui jah, siis tuleb testi p-väärtus väiksem kui 0.05, kui mitte, siis saame p-väärtuse, mis on suurem kui 0.05.*

Lisaks normaaljaotustestile saab normaaljaotuslikkuse testimiseks kasutada ka eelnevalt jutuks olnud asümmeetria- ja järsakuskordajat. Andmed on ideaalselt normaaljaotuslikud kui mõlema statistiku väärtused on võrdsed nulliga. Sellist olukorda tuleb empiiriliste ehk mõõtmise kaudu saadud andmete korral harva ette. Seetõttu peetakse kokkuleppeliselt normaaljaotuslikuks muutujaid, mille asümmeetria ja järsakus jäävad -1 ja 1 vahele. Seegi vahemik ei ole tegelikkuses kivisse raiutud, sest liberaalsemate hinnangute kohaselt võiks selleks vahemikuks olla hoopis -2 ja 2. Kahjuks ei ole siin võimalik öelda, kumba neist lähenemistest tuleks eelistada. Küll aga võib olla kasu soovitusest olla neis valikutes nii läbipaistev kui võimalik. Teiste sõnadega, mis tahes kriteeriumi te ei kasutaks, proovige enda analüüside osa kirjeldamisel olla võimalikult läbipaistev ja kirjeldada, millist kriteeriumi kasutasite.

4. STANDARDISEERITUD SKOORID

Muutuja standardiseerimiseks nimetatakse teisendust, kus muutuja igast väärtusest lahutatakse aritmeetiline keskmine ning saadud vahe jagatakse standardhälbega ehk

$$z_i = \frac{x_i - \bar{x}}{\delta},$$

kus x_i vastab vaatlusele i, $\bar{x}$ aritmeetilisele keskmisele ja δ (sigma) tähistab standardhälvet. Saadud väärtusi nimetatakse ka z-skoorideks. Et muutuja standardiseerimine oleks sisukas peaks standardiseeritava tunnuse jaotus olema normaaljaotuse lähedane.

Ülesanne 5. Standardiseerige muutuja *Informeeritus* ja salvestage see muutujana, mille nimi on *Informeeritus_z*.

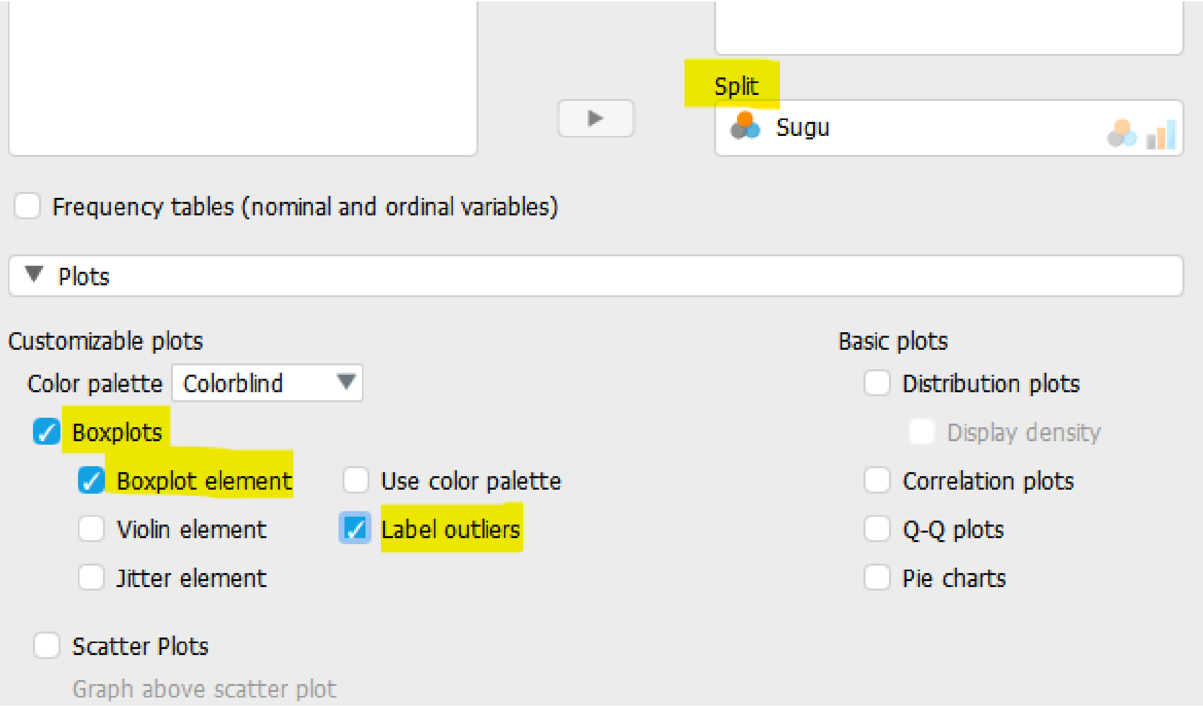
a) Mis on saadud muutuja aritmeetiline keskmine ja standardhälve?
b) Kui tegemist oleks ideaalselt normaaljaotuslike andmetega, siis mitu protsenti vaatlusi jääks -1 ja 1 standardhälbe vahele?

5. HÄLBIVAD VÄÄRTUSED

Ülesanne 6. Joonistage karpdiagramm, millel on naiste ja meeste ruumilise mõtlemise tulemused. Kanna joonisele ka hälbivad väärtused.

a) Mitmendal real paikneb andmetabelis esimene testitäitja, kelle ruumilise mõtlemise tulemus hälbib oluliselt teistest?
b) Kuidas on samal testitäitjal läinud teistes ülesannetes?
c) Kui mitme katseisiku tulemused hälbivad oluliselt ülejäänud testitäitjatest?

Karpdiagrammide joonistamiseks avage kõigepealt *Descriptives* menüü. Tõstke ruumilise mõtlemise alatesti tähistav muutuja *Variables* kasti. Katseisikute soo järgi lahku viimiseks tõsta soo tunnus *Split* kastikesse. Järgmiseks *Plot* menüüst *Boxplots* ning sealt omakorda *Boxplot element*. Selleks, et saada teada, mitmendal real hälbiv väärtus andmetabelis paikneb tehke linnuke *Label outlier* ette, mis kleebib igale hälbivale väärtusele külge rea numbri.



JASP kasutab hälbivate väärtusete tuvastamiseks meile eelmistest ülesannetest tuttavaid protsentiile ja täpsemalt Tukey reeglit. Selle reegli kohaselt peetakse oluliselt hälbivaks neid vaatlusi, mis jäävad 1.5 kvartiilvahemikku allapoole 25-ndast ja ülespoole 75-ndast protsentiilist ehk 1.5 kvartiilvahemikku allapoole esimesest ja ülespoole kolmandast kvartiilist. Hälbivad väärtused märgitakse joonisel haaradest väljapoole jäävate mustade täpikestena.

4. ÜLESANDED ISESEISVAKS HARJUTAMISEKS

Avage fail p2-*InvisibilityCloak.csv* ja lahendage sellel järgmised ülesanded:

Ülesanne 1. Tekitage uus muutuja, mis jagab andmefaili tunnuse *Mischeif* viiekümnenda protsentiili alusel kaheks grupiks.

Ülesanne 2. Tekitage muutujast *Mischeif* uus muutuja *Mischeif* _z, milles on *Mischeif* -i väärtused standardhälbe ühikutes.

Ülesanne 3. Arvuta muutujatele *Mischeif* ja *Mischeif* _z-ile asümmeetria- ja järsakuskordaja. Joonistage ka histogrammid.

Ülesanne 4. Filtreerige andmetabelist välja katseisikud 10 ja 11.

Ülesanne 5. Kas andmestikus on tegemist sõltuvate või sõltumatute gruppidega?

III PRAKTIKUM: KESKMISTE VÕRDLEMINE KAHE GRUPI KORRAL

Järelduste tegemisel ei piisa sellest, kui me näitame, et kahe grupi keskmised on erinevad – lisaks on vaja teada, kas see leid on statistiliselt oluline või mitte. T-testiga saamegi uurida, kas erinevus kahe grupi keskmiste vahel on statistiliselt oluline.

T-testi eeldused

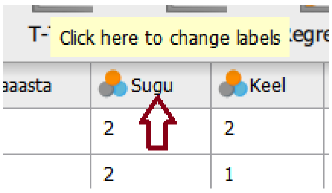
1. Sõltuv muutuja on vähemalt intervallskaalal.
2. Sõltumatu muutuja on kategoriaalne tunnus ehk tunnus, mis jaotab uuritavad kahte gruppi.
3. Sõltuv muutuja peab olema ligilähedane normaaljaotusele:
   a. Sõltuvate rühmade t-testi puhul peab olema normaaljaotusega jaotus, mis on moodustatud järgmiselt: ühel korral mõõdetud tulemustest on lahutatud teisel korral mõõdetud tulemused;
   b. Sõltumatu t-testi puhul peab tunnus rühmade lõikes olema ligikaudu normaaljaotusega.
4. Sõltumatute gruppidega t-testi puhul peaksid olema dispersioonid sarnased (seda saab kontrollida Leveni testiga). Kui gruppide suurused on sarnased, siis ei ole selle eelduse rikkumine väga tõsine. JASPis saab arvutada t-testi statistiku, mis arvestab selle eelduse rikkumisega (Welchi statistik).
5. Sõltumatute gruppidega t-testi puhul peavad valimid olema sõltumatud.

Kui aga eeldused on rikutud (andmed ei ole normaaljaotusega või sõltuv tunnus on järjestusskaalal), siis tuleb kasutada mitteparameetrilist t-testi analoogi. Üldistatult võib öelda, et kui tunnus on arvuline ja ligilähedane normaaljaotusele, saame rakendada parameetrilist statistikat. Järjestustunnustele, normaaljaotusest erinevate jaotuste puhul, väikeste valimite puhul on kohasem rakendada mitteparameetrilist statistikat.

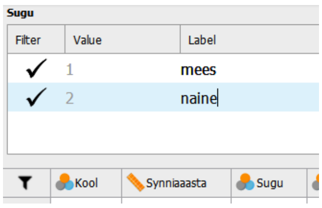
Parameetriline	**Mitteparameetriline**
2 sõltumatu valimiga t test	*Mann-Whitney U Test*
Sõltuva valimiga t test	*Wilcoxon Signed Ranks Test*

1. KAHE SÕLTUMATU GRUPI KESKMISTE VÕRDLEMINE

Kahe sõltumatu grupi keskmiste erinevuse uurimiseks kasutame kahe sõltumatu grupiga t-testi.
Avage andmefail **P1_AKtest.csv** JASP-is.
Enne analüüsiga alustamist vaatame üle andmestiku. Näeme, et muutuja Sugu koosneb väärtustest „1“ ja „2“. Anname neile sildid, mis teevad nende kategooriate lugemise meile lihtsamaks. Vajutage hiirega muutuja nimele:



Meie praegustes andmetes on grupp 1 mehed ja grupp 2 naised. Seega lisame vastavad sildid nende väärtuste juurde:

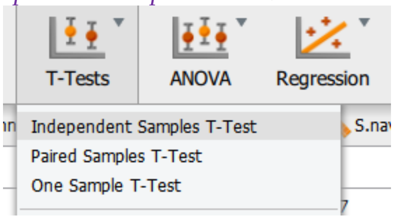


Nüüd näeme andmestikus meie lisatud silte.

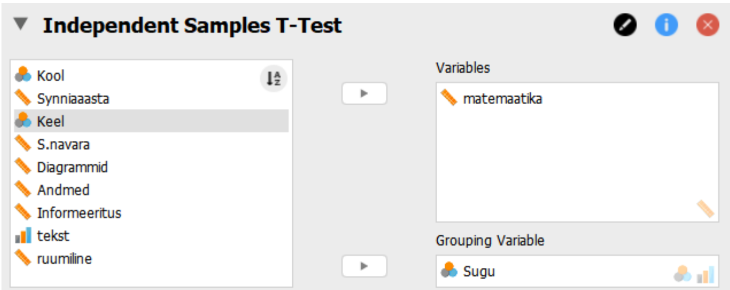
Ülesanne 1. Testime hüpoteesi: meeste matemaatika tulemus on statistiliselt oluliselt kõrgem kui naiste matemaatika tulemus.
Teeme selle analüüsi järk-järgult läbi.

1) Esialgu vaatame, kas meie sõltuv muutuja on normaaljaotusega. Tuletame meelde, et sõltumatute gruppidega t-testi puhul pidime seda tegema kategooriate lõikes (gruppide eraldamiseks lisage grupi tunnus Sugu *Split* aknasse.
Selleks arvutame asümmeetriakordaja (*skewness*) ja järsakusastme (*kurtosis*) näitajad. Kokkuleppeliselt on tegemist normaaljaotusega, kui asümmeetriakordaja ja järsakusastmekordaja väärtused on vahemikus [-2; 2], konservatiivsemalt ka [-1; 1]. Kirjeldava statistika tegemiseks valime ülemisest menüüst *Descriptives - Descriptive Statistics*. Asümmeetriakordaja ja järsakusastme arvutamiseks valime alamenüüst *Statistics skewness* ja *kurtosis*.

2) Kui oleme veendunud, et meie sõltuv muutuja on ligilähedaselt normaaljaotuslik, siis saame rahulikult jätkata t-testi läbiviimisega.
Valime menüüribalt *T-Tests – Independent Samples T-Test*.

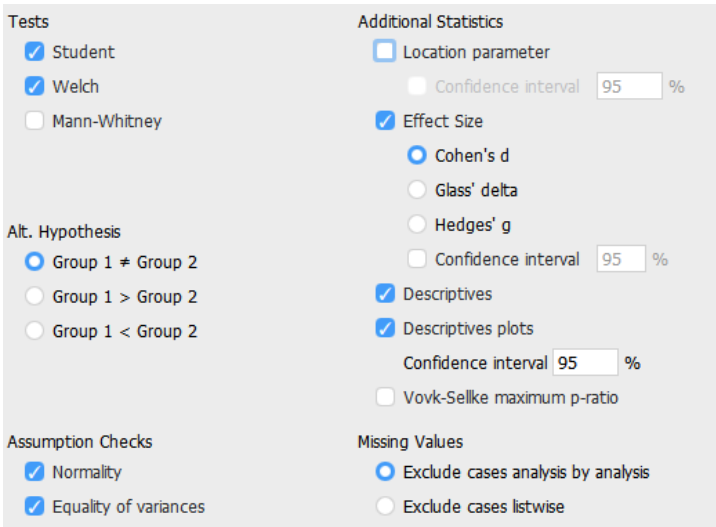


Seejärel avaneb analüüsimenüü *Independent Samples T-Test*. Lisame aknasse *Variables* meie sõltuva muutuja ja aknasse *Grouping Variable* meie sõltumatu muutuja. Sõltuv muutuja on meil seekord matemaatika testi tulemus ja sõltumatu muutuja ehk grupeeriv tunnus on sugu.



Järgnevalt valige allolevast menüüst:

- *Student* ja *Welch* testid. Esimene on tavaline t-test. Welch testi kasutatakse juhul, kui dispersioonide hajuvuse eeldus on rikutud;
- Eelduste testimise (*Assumption checks*) alt *Normality* ja *Equality of variance*. Esimene valik lisab meie väljundisse Shapiro-Wilki testi tulemuse, mis samuti hindab andmete normaaljaotuslikkust. Teine valik viib läbi Leveni testi, mis hindab dispersioonide sarnasust;
- Efekti suuruste puhul valige *Coheni d*;
- Kirjeldav statistika (*Descriptives*);
- Kirjeldavad joonised (*Descriptive plots*).



3) Normaaljaotuse hindamine Shapiro-Wilki testi alusel.

Test of Normality (Shapiro-Wilk)

		W	p
matemaatika	mees	0.975	< .001
	naine	0.962	< .001

Note. Significant results suggest a deviation from normality.

Tabelist tuleb vaadata ainult testi p-väärtust. **Kui p-väärtus on väiksem kui 0.05, siis ei ole Shapiro-Wilki testi alusel andmed normaaljaotuslikud.** *Kas meie ülesandes on andmestik normaaljaotusega?*

Praktikas aga on omajagu harv normaaljaotustestide põhjal leida normaaljaotuslikkust – levinud on asümmeetriakordaja (*skewness*) ning ekstsessi (*kurtosis*) vaatamine. Nii K-S kui ka S-W testidel on omad probleemid; üheks neist on liigne tundlikkus äärmuslikele väärtustele ehk erinditele (*outliers*).
Andmeid peetakse normaaljaotuslikuks siis, kui nii asümmeetriakordaja kui ka järsakusaste/ekstsess on vahemikus (-1;1); liberaalsemalt on aga levinud ka vahemike (-2; 2) kasutamine.

Kuigi me näeme, et Shapiro-Wilki testi alusel ei ole need andmed normaaljaotusega, siis võime siiski võtta liberaalsema seisukoha ja öelda, et asümmeetriakordaja ja järsakusaste olid piisavalt head, et kasutada parameetrilist statistikat. Viimane lähenemine on sotsiaalteadustes üsna tavaline.

4) Dispersioonide hindamine Levene'i testi alusel

Test of Equality of Variances (Levene's)

	F	df	p
matemaatika	31.034	1	< .001

Levene'i testi tulemus annab meile teada kumba rida t-testi tabeli puhul peaksime vaatama (Student t-test on ülemine; Welchi test on alumine). **Kui Levene'i testi Sig on suurem kui 0.05, vaatame edaspidi ülemist tabelirida** (näitab, et jaotuste „kujud" ei erine statistiliselt oluliselt)**; kui Levene'i test Sig on väiksem kui 0.05, loeme edaspidi alumist rida** (näitab, et jaotuste „kujud" erinevad statistiliselt oluliselt).

Kumma rea tulemusi me antud ülesandes vaatama peame?

5) **T-testi tulemus**. Kui Levene'i test näitas, et dispersioonid on sarnased, siis vaadake ülemist rida ehk Student testi tulemust, kui aga Leveni test näitas, et dispersioonid on erinevad, siis vaadake alumist rida ehk Welchi testi tulemust.

Independent Samples T-Test

	Test	Statistic	df	p	Cohen's d
matemaatika	Student	9.253	1348.000	< .001	0.506
	Welch	9.110	1198.429	< .001	0.502

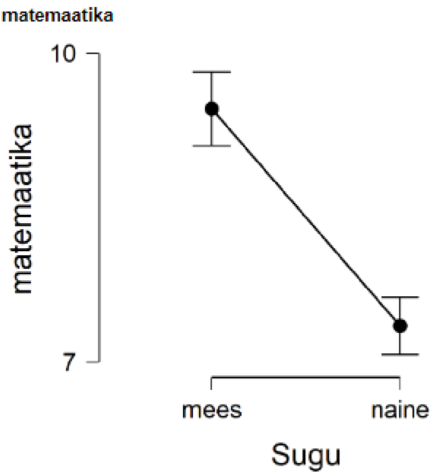
Antud tabelis on kirjas t-statistik, vabadusastmed (df), p-väärtus ja efekti suurus. Statistikas on saanud traditsiooniks kasutada **olulisusnivoosid** 0.01 (ehk 1%) ja 0.05 (ehk 5%). Valides olulisusnivooks 0.05, peab olulisustõenäosus selleks, et nullhüpoteesi ümber lükata, olema väiksem kui 0.05 ning vastavalt olulisusnivoo 0.01 korral peab ta olema väiksem kui 0.01. Seega kui p-väärtus on alla olulisusnivoo, siis on tegemist statistiliselt olulise tulemusega. Selle analüüsi puhul näeme, et tulemus on isegi alla 0.001. See tähendab, et tegemist on väga väikse p-väärtusega. Seega on selle testi tulemus ka statistiliselt oluline. Me saame selle tulemuse põhjal järeldada, et meeste ja naiste tulemused matemaatika testis on erinevad. Efekti suurus[i] on statistiline näitaja, mis võimaldab lisaks statistilisele olulisusele kirjeldada gruppidevahelisi erinevusi. Efekti suurust saab väljendada mitmete statistikutega; ilmselt levinuim on ***Coheni d.*** Kokkuleppeliselt tähistavad Coheni *d*[ii] väärtused väikest efekti väärtusel d = 0.2; keskmise suurusega efekti väärtus on d = 0.5; suure efekti väärtuse algus on d = 0.8.

6) Kirjeldav statistika (*group statistics*) – antud tabelis on ära toodud mõlema grupi suurused, keskmised, standardhälbed ja standardvead. Need aitavad meil neid tulemusi paremini mõista. Samuti peaks tulemuste raporteerimisel ära märkima keskmised ja standardhälbed.

Group Descriptives

	Group	N	Mean	SD	SE
matemaatika	mees	608	9.461	4.516	0.183
	naine	742	7.352	3.856	0.142

Descriptives Plot ▼



7) **Tulemuste raporteerimine** - teadmisest, kuidas näevad välja gruppide keskmised ning kas gruppidevaheline erinevus on oluline või mitte, üksi ei piisa. Need tulemused tuleb kuidagi ka kirjapilti saada. T-testi raporteeritakse järgnevalt[iii]:

*Selles suvalises näidislauses leiti, et loengutes kohalkäijate keskmine tulemus (**M** = 4.51, **SD** = 0.30) on statistiliselt oluliselt kõrgem kui neil, kes magavad sisse ja kohale ei tule (M = 2.92, SD = 0.31), **t(<u>kirjuta siia df väärtus</u>)** = (<u>kirjuta siia t väärtus</u>), **p** = 0.008, **[efektisuuruse statistik]** = (<u>kirjuta efekti suurus siia</u>).*

Kuidas me enda ülesande, mis puudutab meeste ja naiste erinevust matemaatikas, tulemuse kirja paneme?

Mida aga teha kui parameetrilise testi eeldused on rikutud?

Mis juhtub siis, kui parameetriliste testide parameetrid (ehk eeldused siin kontekstis) ei ole täidetud? Väga lihtne – appi saab võtta mitteparameetrilised testid[iv].

Mitteparameetrilise testi tegemiseks valige uuesti menüüribalt *T-Tests – Independent Samples T-Test*. Seekord aga valige avanenud alamenüüst ainult:

- Mann-Whitney
- Effect size

Ülesanne 2. Tehke läbi eelmine ülesanne ka mitteparameetrilise testiga (oletame lihtsalt, et t-testi tegemise eeldused on rikutud). Seega vaatame uuesti, kas meeste ja naiste keskmised tulemused matemaatika testis on erinevad, aga seekord kasutame selleks Mann-Whitney U testi.
Missuguse tulemuse saite? Kas tulemus oli statistiliselt oluline?

Tulemuste raporteerimine mitteparameetrilise testi puhul:

Mann Whitney U test näitas, et naiste (mediaan = ...) ja meeste (mediaan = ...) keskmised tulemused sõnavara testis ei erinenud statistiliselt olulisel määral, U= ... , p = ...

Missugused väärtused peaksime sellesse lausesse sisestama? (vihje: U-statistik on tabelis veerus nimega „W“)

2. KESKMISTE VÕRDLEMINE KAHE *SÕLTUVA* RÜHMA KORRAL

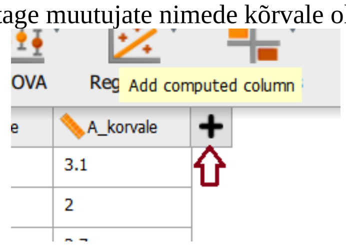
Avage andmefail **AUhinnangud.csv**. Tegemist on ühe koolitöö raames kogutud andmestikuga, milles näidati naisterahvastele 10 mehe pilte, mida paluti atraktiivsuse (A) ja usaldusväärsuse (U) 10-palli-skaalal hinnata. Igast pildist oli 2 versiooni: (1) ühel juhul vaatas isik otse, (2) teisel juhul oli pilti töödeldud nii, et pilk oli suunatud kõrvale. Hüpoteesideks oli, et otsepilguga pilte hinnati keskmiselt (H_1) atraktiivsemaks ning (H_2) usaldusväärsemaks. Kas need hüpoteesid leidsid kinnitust ka andmestikust?

Ülesanne 3. Uurige välja, kas usaldusväärsuse hinnangutes oli erinevus vastavalt sellele, kuhu oli pilk pööratud. Selleks kasutage sõltuvate valimitega t-testi (***Paired Samples t test***).
Miks me siin olukorras <u>sõltuvate</u> valimitega t-testi kasutame?

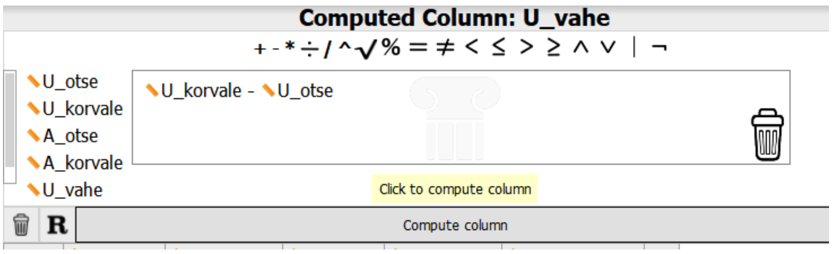
Lahenduskäik on sarnane sõltumatute valimitega t-testi korral; paar olulist erinevust siiski on. Kui sõltumatu t-testi puhul kontrollisime *sõltuvate muutujate* eeldusi, siis sõltuvate valimitega t-testi korral tuleb kontrollida esimese ja teise (või ajaliselt mõne muu) *mõõtmiskorra vahe* normaaljaotuslikkust . Selleks tekitame mõlema hinnangu kohta (usaldusväärsuse ja atraktiivsuse) uued muutujad: usaldusväärsuse hinnangute vahe ja atraktiivsuse hinnangute vahe.

NB! Uue muutuja arvutamine:

1) Uue muutuja arvutamiseks vajutage muutujate nimede kõrvale olevale pluss märgile:



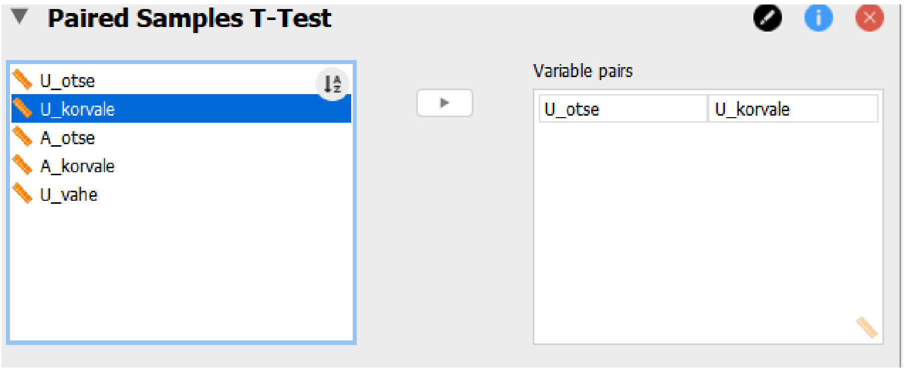
2) Järgnevalt tuleb ekraanile aken, kuhu saate sisestada uue muutuja nime. Paneme selleks „U_vahe“.
3) Andmestiku üleval peaksite nüüd nägeme kasti nimega „Computed Column: U_vahe. Sinna kasti saame lisada uue veeru arvutamiseks vajaliku tehte. Hetkel tahame lahuta ühest usaldusväärsuse hinnangust maha teise usaldusväärsuse hinnangu ehk „U_otse – U_korvale“. Matemaatilise tehte lisamiseks peame selle lohistama kasti ülal olevast nimekirjast alla kasti. Seejärel saame lisada vajalikud muutujad tehtesse. Lõpetuseks vajutage „*Compute column*“ nupule.



Tehke samamoodi ka uus muutuja atraktiivsuse hinnangute alusel. Kui need muutujad on olemas, saate nende normaaljaotuslikkust testida juba harjumuspärasel viisil: arvutage asümmeetriakordajad ja järsakusastmed (*Descriptives – Descriptive Statistics – Statistics: Skewness, Kurtosis*).

Kas muutujad on normaaljaotuslikud?

Sõltuvate rühmadega t-testi läbiviimiseks valige: ***T-tests* – *Paired-Samples T-Test***. ***Variable pairs*** alla tõstke analüüsiks vajalikud muutujate paarid. Ülejäänud valikud on samad, mis tavalise t-testi puhul.



Kuidas te tulemusi tõlgendate? Kas tulemused olid statistiliselt oluliselt erinevad?

Tulemuste raporteerimine:

Sõltuvate rühmadega t-test näitas, et otse vaatava pilguga piltide (M = ...; SD = ...) usaldusväärsuse hinnangud olid statistiliselt oluliselt kõrgemad kui kõrvale vaatava pilguga piltide (M = ...; SD = ...) hinnangud (t (...) =...., p<. ...., d = ...).

Ülesanne 4. Uurige välja, kas atraktiivsuse hinnangutes oli erinevus vastavalt sellele, kuhu oli pilk pööratud.

MITTEPARAMEETRILISE TESTIGA

Kui te aga leiate, et andmed ei ole normaaljaotuslikud või on tegemist hoopis järjestusskaalal kogutud andmetega, siis tuleks kasutada mitteparameetrilist testi, sõltuvate rühmadega t-testi puhul on selleks Wilcoxon's signed-rank test.
Mitteparameetrilise testi tegemiseks valige uuesti menüüribalt *T-Tests – Paired Samples T-Test*. Seekord aga valige avanenud alamenüüst ainult:

- Wilcoxon signed-rank
- Effect size

Ülesanne 5. Proovige arvutada ka usaldusväärsusele ja atraktiivsusele antud hinnaguid mitteparameetriliste testidega.

Mitteparameetrilises analoogis on olulisel kohal *Wilcoxoni Z* (sõltumatute valimite puhul *Mann-Whitney U*).

Need statistikud tuleb teil raporteerida järgnevalt:

Wilcoxoni Signed Ranks Test näitas, et otse vaatava pilguga pilte (mediaan = ...) hinnati statistiliselt oluliselt atraktiivsemaks kui kõrvale vaatava pilguga pilte (mediaan = ...), Z = ..., p = ...

Alternatiivselt:

Wilcoxoni Signed Ranks Test näitas, et otse vaatava pilguga piltide (mediaan = ...) atraktiivsuse astakud olid statistiliselt oluliselt kõrgemad kui kõrvale vaatava pilguga piltide (mediaan = ...) astakud, Z = ..., p = ...

JASP-is on t-testide analüüside aknas saadaval veel üks valik – One Sample t-test. Seda me antud kursuse raames rohkem ei käsitle, kuid hea on teada, et sellega on võimalik teha t-testi olukorras, kus tahame teada, kas meie andmed erinevad mingist etteantud väärtusest. Sellest võib mõelda kui t-testist kahe grupi vahel, kus ühe grupi andmete asemel on meil olemas ainult üks väärtus – selle grupi keskmine. Teisisõnu võrdleme oma andmeid (ühe grupi skoorid või näiteks kahe grupi skooride vahe) ühe etteantud fikseeritud väärtusega (antud väärtuse saab sisestada Tests all leitavas aknas nimega Test Value.
See analüüs võib osutuda kasulikuks juhul, kui a) teame populatsiooni tegelikku või arvatavat keskmist ning tahame teada, kas meie valim on sellest oluliselt erinev, b) teame mingi testi keskmist võimalikku skoori ning tahame teada, kas meie valimi testitulemused on sellest oluliselt erinevad, c) tahame võrrelda oma valimi tulemusi juhuslikult saadava tulemusega (selleks peab eelnevalt välja arvutama, mis väärtuse puhul oleks tulemus juhuslik), d) tahame võrrelda oma valimi tulemusi (või kahe grupi tulemuste erinevust) nulliga. NB Antud test eeldab samuti, et tegemist on sõltumatute gruppidega!

3. KOLME JA ROHKEMA GRUPI KESKMISTE VÕRDLEMINE

Sissejuhatus ANOVA-sse (jätkub järgmises praktikumis)

Kui võrreldavaid gruppe on rohkem kui kaks (seega 3 või enam), siis tuleb kasutada dispersioonanalüüsi ehk ANOVAt (***An**alysis **o**f **v**ariance*). ANOVA analüüse on mitmeid, sõltuvalt vajadusest. Neid saab samuti jagada laias laastus selle alusel, kas grupitunnused on mõõdetud erinevate (sõltumatud grupid) või samade inimeste peal (sõltuvad grupid).

1) Sõltumatud grupid (*between subjects ANOVA*):

   Ühefaktoriline dispersioonanalüüs ehk *One-way ANOVA** - Kui meil on üks arvuline sõltuv muutuja ja üks kategooriline sõltumatu muutuja (faktor), millel on rohkem kui kaks taset, siis saab kasutada ühefaktorilist ANOVAt.

   Näiteks uuritakse, kas keskmine sissetulek (sõltuv muutuja) erineb keskharidusega, bakalaureusekraadiga ja magistrikraadiga (I faktor, 3 taset) inimeste vahel.

   Faktoriaalne dispersioonanalüüs ehk *Factorial ANOVA** (ka *Two-way ANOVA* ja *Three-way ANOVA*) – Kui meil on üks arvuline sõltuv muutuja, aga mitu kategoorilist sõltumatut muutujat (tavaliselt 2 või 3), millel on omakorda mitu taset, siis saab kasutada faktoriaalset ANOVAt.

   Näiteks uuritakse, kas keskmine sissetulek sõltub lisaks haridustasemele (I faktor, 3 taset) ka soost (II faktor, 2 taset)

   *Neid analüüse saab JASPis teha valides ülemisest menüüst *ANOVA – ANOVA, kus sõltuv muutuja tuleb lohistada kasti Dependent Variable ning grupi tunnus kasti Fixed Factors*

2) Sõltuvad grupid (*within subjects ANOVA*):

   Korduvmõõtmiste dispersioonanalüüs ehk *Repeated Measures ANOVA*** – Kasutatakse siis, kui sõltumatud muutujad on mõõdetud samal valimil. Ka siin võib faktoreid olla üks või mitu.

   Näiteks uuritakse, kas inimeste sooritus tähelepanuülesandes (sõltuv muutuja) on erinev sõltuvalt sellest, kas talle mängitakse kõrvaklappidest klaverimuusikat, rokk-muusikat või valget müra (I faktor, 3 taset) - iga inimene teeb ülesanded läbi kolm korda erineva taustamuusikaga.

3) Segatüüpi dispersioonanalüüs ehk *Mixed-model ANOVA*** – korduvmõõtmiste ANOVA, millel on lisaks sõltuvate gruppide faktorile (*within-subjects*) ka mõni sõltumatute gruppide faktor (*between-subjects*)
   Näiteks uuritakse, kas lisaks taustamuusikale (I faktor, 3 taset, sõltuvad grupid) võiks sooritust tähelepanuülesandes mõjutada ka katseisiku sugu (II faktor, 2 taset, sõltumatud grupid).

**Neid analüüse saab JASPis teha valides *ANOVA – Repeated Measures ANOVA*

Dispersioonanalüüsi eeliseks on võimalus mõõta mitme sõltumatu muutuja efekti korraga – see tähendab seda, et saame eristada nii faktorite **peamõjusid** kui ka nendevahelisi **interaktsioone (koosmõju).** Statistiliselt oluline interaktsioon ütleb meile seda, et ühe faktori efekt sõltub teise faktori tasemest.

4. ÜLESANDED ISESEISVAKS HARJUTAMISEKS:

Kasutage andmefaili P1_AKtest.csv.

Ülesanne 1. Uurige hüpoteesi: naiste teksti mõistmise testi tulemus on statistiliselt oluliselt parem kui meeste oma.

Kas naiste ja meeste dispersioonid erinevad selle testi lõikes?Mis on selle väite olulisuse tõenäosus?
Tehke ka joonis, kus on keskmised testi tulemused koos usalduspiiridega.

Ülesanne 2. Uurige hüpoteesi: meeste ja naiste keskmised tulemused informeerituse testis on erinevad.
Mis on selle tulemuse efektisuurus? Kas see efekt on väike, keskmine või suur?

Ülesanne 3. Tekitage uus muutuja nimega „AKskoor“, liites kokku kõikide 7 testi alaskaala tulemused. Kas kahe kooli vahel on statistiliselt oluline erinevus selles tulemuste koondskooris?
Mis on selle küsimuse puhul alternatiivne hüpotees ja mis nullhüpotees?

Ülesanne 4. (Kordamine) Tekitage uus muutuja nimega „Vanus_grupp“ kasutades selleks ifElse avaldist, nii et 90-92.a sündinud inimesed moodustavad grupi „1“, 85-89.a sündinud grupi „2“ ning 80-84.a grupi „3“ (lepime sellega, et grupid tulevad ebavõrdsed). *NB Vihje, kuidas sellist avaldist kirjutada, leiab 2.praktikumi juhendist.*
Mis tüüpi skaalaga on selle uue muutuja puhul tegu?

Ülesanne 5. Proovige teha ühesuunaline ANOVA, kus sõltumatuks muutujaks on vanusegrupp. Tellige usalduspiiridega joonis - kas sõnavaratesti skoor võiks olla erinevate vanusegruppide puhul erinev? Aga alatestide koondskoor AKskoor? Milline vanusegrupp joonise järgi teistest erineb, kui üldse?

Lisaülesanne: Mõelge põnev või naljakas näide, mille uurimiseks oleks vaja kasutada: a) ühesuunalist ANOVAt, b) faktoriaalset ANOVAt, c) korduvmõõtmiste ANOVAT. Tooge välja ka nende faktorite tasemed.

---

[i] http://www.leeds.ac.uk/educol/documents/00002182.htm - arusaadavat ja lihtsasti loetavat lisamaterjali efekti suurusest.
[ii] http://rpsychologist.com/d3/cohend/ - Cohen-i d tõlgendamisest.
[iii] http://goo.gl/jLKFqN - kuidas APA-stiilis tulemusi raporteerida
[iv] http://www.stats.gla.ac.uk/steps/glossary/nonparametric.html - lühikokkuvõte mitteparameetrilistest testidest keskmiste võrdlemisel.

IV PRAKTIKUM: KESKMISTE VÕRDLEMINE ENAM KUI KAHE GRUPI KORRAL

Selles praktikumis vaatleme dispersioonanalüüsi erinevaid liike: a) One-Way ANOVA-t ehk ühefaktorilist ANOVA-t, b) mitme faktoriga ANOVA-t ehk faktoriaalset ANOVA-t ja c) korduvmõõtmiste ANOVA-t, d) õpime leidma olulisuse tõenäosust, efekti suurust arvutama ja tõlgendama ning post hoc analüüse teostama, e) korduvmõõtmiste puhul õpime kontrollima sfäärilisuse (*sphericity*) eeldust ja f) eelduse rikkumisel korrigeeritud *p*-väärtust ja vabadusastmete arvu raporteerima.

Sageli huvitab uurijaid enam kui kaks gruppi. Näiteks võib ravimiuurijaid huvitada, kas ravimi toime erineb platseebost, kuid ka näiteks, kas selle mõju on parem või võrreldav hetkel turul oleva parima alternatiiviga. Sellisel juhul on sõltuvaks muutujaks (ehk muutujaks, milles uurijad loodavad erinevust näha) sümptomite hulk ja sõltumatuks muutujaks (ehk mida eksperimentaator süstemaatiliselt varieerib ja millel on eeldatavalt mõju sümptomite hulgale) ravimi tüüp, millel on kolm taset: 1) platseebo, 2) uus ravim ja 3) vana ravim. Dispersioonanalüüs ongi mõeldud täpselt sellisteks puhkudeks, kus tahame mitme grupi keskmisi võrrelda. Ravimiuuringud on ilmselt üks lihtsamini näitlikustatav valdkond dispersioonanalüüsi ehk ANOVA (*Analysis Of Variance*) kasutamisest. Aga meetod on laialt kasutatav väga paljudes teistes mitut gruppi (st enamat kui kahte gruppi) võrdlevates uuringutes. Selles praktikumis vaatamegi mõningaid juhtumeid, kus sobiks gruppide võrdlemiseks kasutada dispersioonanalüüsi.

Tegelikult varjab ANOVA nime taga terve perekond erinevaid analüüse, mis võrdlevad uurija poolt varieeritud kategooriate mõju uuritavale pidevale tunnusele. Näiteks eelmises näites sobiks kasutada ühesuunalist (One-Way) ANOVA-t, sest uuritakse ühe kategoriaalse sõltumatu muutuja mõju. Tihti huvitab uurijaid aga kahe või enama kategoriaalse sõltumatu muutuja panus ja sellisel juhul kasutatakse faktoriaaleset ANOVA-t. Kahe sõltumatu muutuja puhul kutsutakse seda ka Two-Way ANOVA-ks. Eelnevalt toodud näites oli tegemist sõltumatute gruppidega, sest mõõdeti erinevaid inimesi, kellele manustati üht kolmest ravimi tüübist. Väga paljudes olukordades eelistavad aga uurijad sellele lähenemisele sõltuvate gruppide katsedisaini ehk samade inimeste reaktsioonide mõõtmist erinevates tingimustes. Üheks motivaatoriks on see, et inimeste vaheline variatiivsus on pahatihti suurem kui see variatiivsus, mida suudab meie sõltumatu muutuja andmetes kirjeldada. Sõltuvate gruppide disaini kasutades hoiame aga individuaalsetest erinevustest tingitud variatiivsuse kontrolli all sellega, et vaatleme samu indiviide samades tingimustes. Sellist lähenmist kutsutakse korduvmõõtmiste disainiks ja nende andmete analüüsiks sobib ANOVA perekonnast korduvmõõtmiste ehk *repeated measures* ANOVA. Kuigi erinevad nimetused võivad esialgu segadust tekitada, siis lohutuseks on teadmine, et tegelikult on tegemist samasse perekonda kuuluvate testidega, mille tõlgendamine on väga sarnane.

1. ÜHEFAKTORILINE ANOVA

Avage andmefail **p4_clinicaltrial.csv**. Tegemist on simuleeritud ravimiuuringute andmestikuga; muutujateks on katseisiku number (ID), ravimi tüüp (platseebo, anxifree ja joyzepam), info teraapias osalemise kohta (*therapy*) ja meeleolu hinnangu muutus pärast sekkumist (mood.gain).

Sõltumatute gruppide ANOVA läbiviimseks peab jälgima sarnaseid eeldusi, mida vaatasime sõltumatu t-testi puhul:

1. Sõltuv muutuja on pideval skaalal ja sõltumatu muutuja on kategoriaalne;
2. Vaatlused on teineteisest sõltumatud;
3. Sõltuv muutuja on gruppide lõikes normaaljaotuslik;
4. Hajuvused on sarnased (Leveni test).

Ülesanne 1. Uurige, kas sõltumatu muutuja erinevate tasemete (platseebo, anxifree ja joyzepam) lõikes on meeleolu hinnangus (mood.gain) statistiliselt olulisi erinevusi?

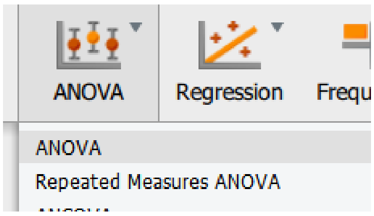
Kui meil on üks sõltuv muutuja ja üks sõltumatu muutuja, millel on rohkem kui kaks taset, siis saame kasutada ühefaktorilist ANOVA-t.

Teeme selle analüüsi järk-järgult läbi.

1) Esialgu vaatame, kas meie sõltuv muutuja on normaaljaotusega. Sarnaselt sõltumatute gruppide t-testiga hindame seda kategooriate lõikes.

Selleks arvutame asümmeetriakordaja (*skewness*) ja järsakusastme (*kurtosis*) näitajad iga sõltumatu muutuja taseme kohta eraldi (kokkuleppeliselt on tegemist normaaljaotusega, kui asümmeetriakordaja ja järsakusastmekordaja väärtused on vahemikus [-2; 2], konservatiivsemalt ka [-1; 1]). Kirjeldava statistika tegemiseks valime ülemisest menüüst ***Descriptives - Descriptive Statistics***. Asümmeetriakordaja ja järsakusastme arvutamiseks valime ***Statistics*** alamenüüst ***skewness*** ja ***kurtosis***.

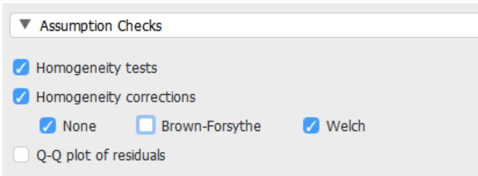
2) Kui oleme veendunud, et meie sõltuv muutuja on ligilähedaselt normaaljaotuslik, siis saame valida menüüribalt ***ANOVA – ANOVA***.



Seejärel avaneb analüüsimenüü *ANOVA*. Lisame aknasse ***Dependent Variable*** meie sõltuva muutuja (mood.gain) ja aknasse ***Fixed Factors*** meie sõltumatu muutuja (drug).

3) Dispersioonide hindamine Levene'i testi alusel. Valige alamenüüst ***Assumption Checks***:

- ***homogeneity tests*** - viib läbi Levene'i testi;
- ***homogeneity corrections***, ***None***, ***Welch*** - lisab ANOVA tabelisse tulemuse, mida peaks vaatama juhul kui Levene'i test näitab, et dispersioonid ei ole sarnased.



Levene'i testi tulemus annab meile teada kumba rida ANOVA tabeli puhul peaksime vaatama (None ehk tavaline ANOVA tulemus on ülemine; Welchi test on alumine). **Kui Levene'i testi *p*-väärtus on suurem kui 0.05, vaatame edaspidi ülemist tabelirida** (näitab, et jaotuste „kujud" ei erine statistiliselt oluliselt)**; kui Levene'i test *p*-väärtus on väiksem kui 0.05, loeme edaspidi alumist rida** (näitab, et jaotuste „kujud" erinevad statistiliselt oluliselt).

Test for Equality of Variances (Levene's) ▼

F	df1	df2	p
1.450	2.000	15.000	0.266

Kumma rea tulemusi me antud ülesandes vaatama peame?

4) **ANOVA tulemus**. Kui Levene'i test näitas, et dispersioonid on sarnased, siis vaadake ülemist rida, kui aga Leveni test näitas, et dispersioonid on erinevad, siis vaadake alumist rida ehk Welchi testi tulemust.

ANOVA - mood.gain

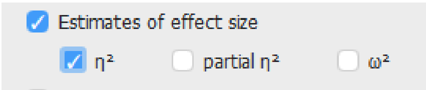
Homogeneity Correction	Cases	Sum of Squares	df	Mean Square	F	p
None	drug	3.453	2.000	1.727	18.611	< .001
	Residuals	1.392	15.000	0.093		
Welch	drug	3.453	2.000	1.727	26.322	< .001
	Residuals	1.392	9.493	0.147		

Note. Type III Sum of Squares

ANOVA tabel näitab meid kõige enam huvitava küsimuse vastust – kas gruppide vahel on statistiliselt olulised erinevused või mitte. ***Veerg p*** (ehk olulisuse tõenäosus) annab meile vastuse. Kui $p < .05$, on meie mudel statistiliselt oluline ning edasi tasub minna juba ***post hoc*** analüüsidega, et välja selgitada, millised tingimused täpsemalt omavahel erinesid.

5) **ANOVA tulemuse efekti suurus.**

Efekti suuruse arvutamiseks valige ***Display*** menüüst ***Estimates of effect size*** ning esimene efekti suuruse statistik ehk tavaline eta ruut (η^2). Seejärel lisab programm automaatselt ANOVA tabelile juurde efekti suuruse. Eta ruut varieerub nullist üheni ning peegeldab kui mitu protsenti sõltuva muutuja hajuvusest meie poolt vaadeldud sõltumatu muutuja kirjeldab (0 vastab nullile protsendile ja 1 sajale).



Miks ei tohi ANOVA asemel teha lihtsalt mitut t-testi? Kui rühmi on rohkem kui kaks peame kasutama ANOVA-t, sest vastasel juhul suurendame tõenäosust, et leiame vale-positiivseid tulemusi (I tüüpi viga suureneb). Näiteks kui teeme kolm võrdlust, siis ei ole meie I tüüpi vea suurus enam 5% vaid 14.3% ($1-(.95)^3$). ANOVA kaasab ühte analüüsi kõik grupid ja annab meile teada, kas meie valimi hajuvus sõltub süstemaatiliselt sõltumatu muutuja(te) erinevatest tasemetest. Teiste sõnadega saame lihtsalt teada, kas gruppide vahel on erinevusi, aga mitte seda, kus see erinevus täpselt paikneb. Kuigi ANOVA akronüüm tuleneb inglise keelsest terminist *Analysis Of Variance*, siis pisikese memotehnilise võttena võib meelde jätta ka lause Ainult Natukene On Võimalik Aimata, sest kui ANOVA tulemus on statistiliselt oluline, siis peame tegema veel lisaks post hoc analüüsi, mis näitab täpsemalt, millised grupid üksteisest oluliselt erinesid. Need võrdlused võtavad arvesse, et suurema hulga võrdluste tegemisel suureneb ka oht saada valepositiivseid tulemusi. Üks lähenemisi, mida kutsutakse bonferoniks korrutab saadud *p*-väärtused läbi võrdluste arvuga. Kuigi selline lähenemine on lihtne, siis praktikas on see liiga konservatiivne ja suurendab II tüüpi vea ohtu (ehk seda, et ütleme, et meie grupid ei erinenud üksteisest, aga tegelikult erinevad). Kursuse õpik soovitab alternatiivina kasutada holmi korrektsiooni, mis on sama hea I tüüpi vigade vältimisel ja edukam II tüüpi vigade vältimisel (Navarro, Foxcroft, & Faulkenberry, 2019, lk 310-311).

6) **Post hoc testid**. Saime ANOVA tabelist teada, et meie mudel oli statistiliselt oluline, aga ANOVA tulemus iseenesest ei ütle meile veel, missugused grupid omavahel erinevad. Selleks peame tegema post hoc testid.

Post hoc testide tegemiseks valige menüüst ***Post Hoc Tests***: ***Holm*** ja ***Standard***.
Post hoc testide tabel näitab erinevate gruppide vahelisi võrdlusi. Oluline on siingi jälgida, millised nendest võrdlustest on statistiliselt olulised.

Post Hoc Comparisons - drug

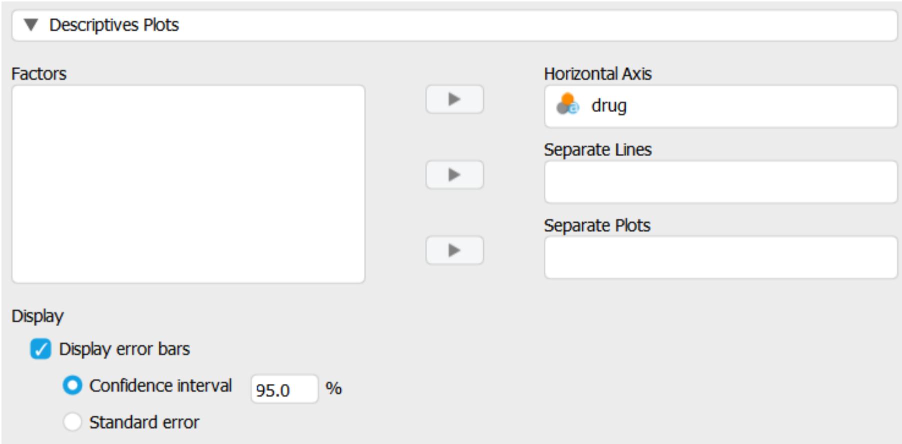
		Mean Difference	SE	t	p_{holm}
anxifree	joyzepam	-0.767	0.176	-4.360	0.001
	placebo	0.267	0.176	1.516	0.150
joyzepam	placebo	1.033	0.176	5.876	< .001

Note. P-value adjusted for comparing a family of 3

7) **Tulemused saab kirja panna järgmiselt:**

Ühesuunaline dispersioonanalüüs (One-Way ANOVA) näitas, et ärevuse hinnang erines oluliselt gruppide vahel ($F(2,15) = 18,6$; $p < .01$, $\eta^2 = 0,71$). Post-hoc keskmiste võrdlused näitasid, et joyzepami manustanud uuritavate keskmine raporteeritud meeleolu oli kõrgeim (M = 1,48; SD = 0,21), erinedes statistiliselt nii platseebost ($p < 0,01$; M = 0,45; SD = 0,28) kui anxifreest ($p < 0,001$; *M* = 0,71; SD = 0,39). Anxifree ja platseebo vaheline võrdlus ei osutunud statistiliselt oluliseks ($p = 0,150$).

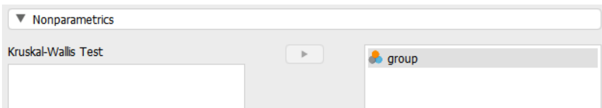
8) **Usalduspiiridega joonis**. Kui tulemus on statistiliselt oluline, siis on kasulik lisada ka illustreeriv joonis. Selleks otsige üles analüüsiaknast menüü ***Descriptive Plots*** ja pange sõltumatu muutuja (antud juhul ravim) x-teljele (***horizontal axis***). Lisage ka joonisele usalduspiirid valides ***Display error bars***.



Meenutuseks 95% usaldusvahemike arvutamiseks kasutati valemit: aritmeetiline keskmine ± (1.96) x (standard viga).

Ülesanne 2. Tehke eelmine harjutus läbi ka mitteparameetrilise testiga. Seega vaatame uuesti, kas meeleolu hinnang (mood.gain) sõltus oluliselt ravimi tüübist (platseebo, anxifree ja joyzepam), aga seekord kasutame selleks Kruskal Wallise testi. Missuguse tulemuse saite? Kas tulemus oli statistiliselt oluline? Mida näitavad post hoc testid?

Kui parameetrilise testi eeldused on rikutud, siis tuleks valida ANOVA analüüsiaknast ***Nonparametrics*** ja viige parempoolsesse aknasse sõltumatu muutuja, mille põhjal tahate mitteparameetrilist testi (antud juhul on selleks Kruskal Wallise test) arvutada:



Post hoc testide tegemiseks avage alamenüüst ***Post Hoc Tests*** ja valige tüüpide alt ***Dunn'i*** post hoc test, mis arvestab sellega, et meie andmed ei ole parameetrilised.

2. FAKTORIAALNE ANOVA (ehk enam kui ühe faktroriga ANOVA)

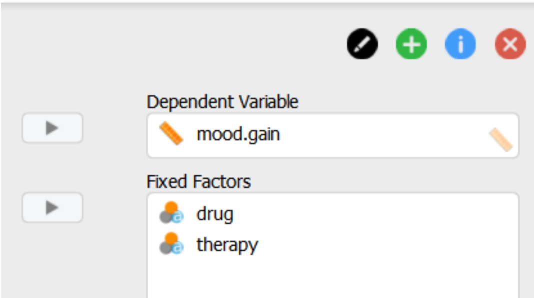
Eelnevat ülesannet täites nägime, et lisaks meie poolt juba kasutatud sõltumatule muutujale (*drug*) oli andmetabelis veel üks kategoriaalne tunnus (*therapy*), mis tähistab seda, kas uuringus osalejad käisid uuringu toimumise ajal ka kognitiivkäitumuslikus teraapias. Kuna reaalses elus määratakse ravimeid tihtipeale kombinatsioonis mõne teraapiaga, siis võivad uurijad (ja ravimeid välja kirjutavad arstid ning nende patsiendid) tahta teada, kas ravim töötab erinevalt (näiteks efektiivsemalt) neil, kes samal ajal ka teraapias käivad. Eksperimentaatorid jagasid katsealused gruppidesse selliselt, et igasse ravimi gruppi (platseebo, anxifree ja joyzepam) sattus sama palju inimesi, kes käisid teraapias ja kui neid, kes ei käinud. Sellise katse disainist kõnelemisel võidakse öelda, et tegu oli 3 (platseebo, anxifree ja joyzepam) x 2 (teraapia, mitte teraapia) eksperimendiga, milles sõltuvaks muutujaks oli uuringus osalejate meeleolu.

Seda tüüpi katseplaani puhul räägitakse ka **peaefektidest** ja **interaktsioonist (koosmõjust)**. Eelnevalt toodud näite puhul tähendaks peaefekt näiteks, et uuringus osalejate raporteeritud meeleolu sõltus oluliselt ravimi tüübist (sõltumata sellest, kas teraapias osaleti või mitte). Sellisel juhul oleks tegemist ravimitüübi peaefektiga. Võib olla aga ka nii, et uuringus osalejate meeleolu hinnangud sõltusid teraapias käimisest (sõltumata sellest, kas ravimit võeti või mitte). Võib olla ka nii, et kummalgi faktoril oli oma sõltumatu panus meeleolu hinnangule. Sellisel juhul räägiksime kahest peaefektist (ravimi ja teraapia omast). Neljas võimalus on, et ühe faktori mõju sõltub teisest faktorist. Antud näite kontekstis näiteks, et ravimi efektiivsus sõltub sellest, kas uuringus osalejad käisid samal ajal ka teraapias. Kõigi taoliste faktoriaalsete disainide puhul saame sisuliselt kasutada sama lahenduskäiku nagu ühefaktorilise ANOVA puhul. Peamine erinevus on see, et nüüd teeme mudeli, kus on ühe sõltumatu muutuja asemel kaks sõltumatut muutujat.

Faktoriaalse ANOVA puhul kontrollime samu eeldusi nagu ühesuunalise (One-Way) ANOVA-gi puhul, sest võtame siingi eelduseks, et a) sõltuv muutuja on pideval skaalal ja sõltumatud muutujad on kategoriaalsed, et b) sõltuv muutuja on gruppide lõikes enam-vähem normaaljaotuslik, ning et c) hajuvused on sarnased (Leveni test) ja d) vaatlused üksteisest sõltumatud.

Ülesanne 3. Uurige, kas ravimi ja teraapia peaefektid on olulisad ja kas nende vahel esineb oluline interaktsioon.

Faktoriaalse ANOVA puhul peame lisame aknasse ***Dependent Variable*** meie sõltuva muutuja (*mood.gain*) ja aknasse ***Fixed Factors*** meie kaks sõltumatut muutujat (*drug*, *therapy*).



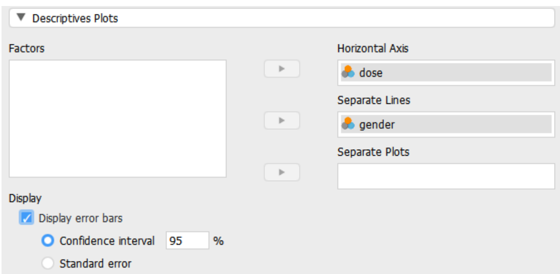
ANOVA tabel annab infot selle kohta, kas mudel, sõltumatute muutujate peamõjud ning interaktsioon on statistiliselt olulised. Antud näite puhul näeme, et nii ravimi kui teraapia peamõjud on statistiliselt oluliselt, kuid ravimi ja teraapia interaktsioon (koosmõju) mitte.

ANOVA - mood.gain

Cases	Sum of Squares	df	Mean Square	F	p	η²
drug	3.453	2	1.727	31.714	< .001	0.713
therapy	0.467	1	0.467	8.582	0.013	0.096
drug ✻ therapy	0.271	2	0.136	2.490	0.125	0.056
Residuals	0.653	12	0.054			

Note. Type III Sum of Squares

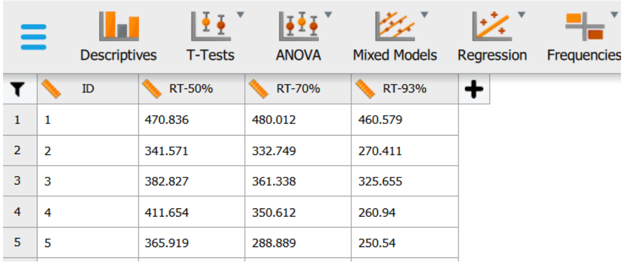
Et tulemustest paremini aru saada tehke ka joonis, kus on mõlemad uuritavad faktorid ära toodud. Üks faktor pange x-teljele ja teine lisage eraldi joontena joonisele:



3. KORDUVMÕÕTMISTE ANOVA (ehk analüüsid enam kui kahe sõltuva grupiga)

Korduvmõõtmiste ANOVA (*Repeated measures ANOVA*) on sarnane tavalise ANOVA-ga (*One Way ANOVA*). Tavalise ANOVA-ga võrdlesime erinevusi sõltumatute gruppide vahel. Korduvmõõtmiste ANOVA-ga saame võrrelda erinevusi sõltuvate gruppide korral.

Avage uus andmefail: **p4_rtData.txt**

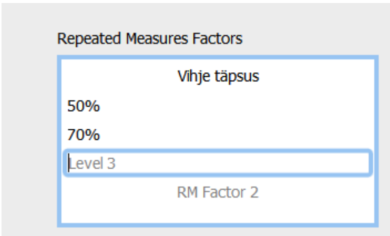


	ID	RT-50%	RT-70%	RT-93%
1	1	470.836	480.012	460.579
2	2	341.571	332.749	270.411
3	3	382.827	361.338	325.655
4	4	411.654	350.612	260.94
5	5	365.919	288.889	250.54

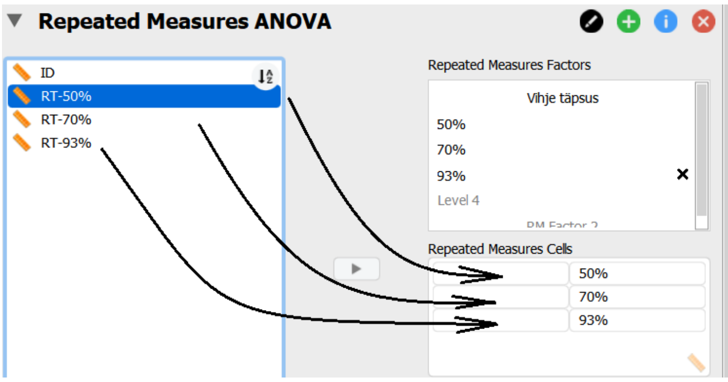
Need andmed pärinevad reaktsiooniaja eksperimendist. Selles katses pidi katseisik vastama võimalikult kiiresti, kas ekraanile esitatud nool näitab vasakule või paremale. Enne noole ilmumist tekkis ekraanile vihje, mis ütles kummale poole nool järgmisena näitab. Kokku oli katses kolme tüüpi vihjeid. Ühed, mis ennustasid peaaegu alati õigesti (93% juhtudest) kummale poole nool järgmisena näitab. Teised, mis ennustasid enamasti õigesti (70%) ja kolmandad, mis ei aidanud katseisikul ennustada kummale poole nool järgmisena näitab (50%).

Ülesanne 4. Vii läbi korduvmõõtmiste ANOVA, et selgitada, kas katseisikute vastamiskiirus sõltus oluliselt vihje täpsusest.

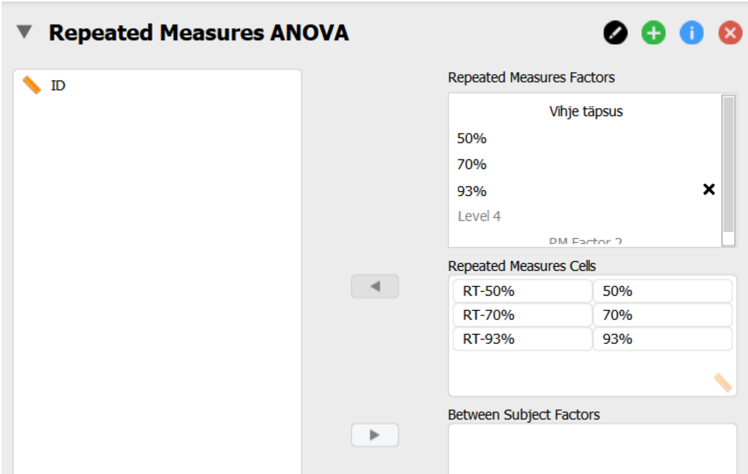
Üldjoontes on lahenduskäik sarnane ühefaktorilise ANOVA-ga. Seekord peame aga faktortasemed eelnevalt käsitsi defineerima. Avage analüüsiaken ANOVA – Repated Measures ANOVA. **Repeated Measures Facotors** aknasse tuleks esmalt kirjutada uuritava faktori nimi ja defineerida selle faktori tasemed. Antud juhul on meil üks faktor. Paneme selle faktori nimeks „Vihje täpsus“ ja kirjutame tasemele 1, 2 ja 3 nimed: 50%, 70%, 93%.



Järgnevalt peaksime aknas **Repeated Measures Cells** viima kokku meie andmestikus olevad muutujad vastavate faktortasemete nimedega (kui muutujad on samas järjekorras nagu meie faktori tasemed, siis võime kõik muutujad ka korraga **Repeated Measures Cells** viia).



Repeated Measures ANOVA sisestusaknad näevad lõpuks välja sellised:



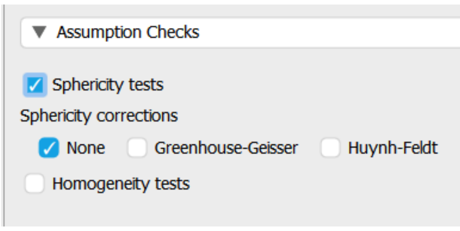
Ülejäänud sammud on sarnased ühefaktorilise ANOVA-ga. Siiski tuleks korduvmõõtmiste puhul vaadata, et erinevate faktortasemete vahede hajuvused oleksid sarnased. Seda eeldust kutsutakse andmete sfäärilisuse (*sphericity*) eelduseks. Õnneks JASP ka teavitab meid selle eelduse rikkumise eest. Sellisel juhul ilmub väljundi aknasse meie tabeli alla märge, et Mauchly test oli positiivne ja seega võib arvata, et eeldus on rikutud (vt alumine joonis).

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p
Vihje täpsus	80215.524[a]	2[a]	40107.762[a]	69.206[a]	< .001[a]
Residuals	44045.227	76	579.542		

Note. Type III Sum of Squares
[a] Mauchly's test of sphericity indicates that the assumption of sphericity is violated (p < .05).

Mauchly testi tellimiseks valige menüüst Assumption Checks ja seal tehke linnuke Sphericity tests ette.

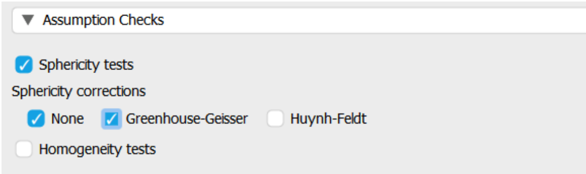


Kui Mauchly testi *p*-väärtus on väiksem kui 0.05, siis on eeldus rikutud ja seega peaksime kasutama F-statistiku raporteerimisel kohandatud vabadusastmete väärtusi ja *p*-väärtust. Kui Greenhouse-Geisseri väärtus on väiksem kui 0.75, siis kasutatakse kokkuleppeliselt Greenhouse-Geisseri korrektsiooni. Vastasel korral Huynh-Feldti nimelist korrektsiooni.

Test of Sphericity

	Mauchly's W	Approx. X²	dfSphericity	p-value	Greenhouse-Geisser ε	Huynh-Feldt ε	Lower Bound ε
Vihje täpsus	0.506	25.215	2	< .001	0.669	0.685	0.500

Mõlema korrektsiooni saab tellida **Assumption Checks** menüüst vastavalt **Greenhouse-Geisseri** või **Huynh-Feldti** ette linnukese tehes.



Seepeale ilmuvad ANOVA tabelisse uued read, mis tähistavad eelduse rikkumise tõttu korrigeeritud vabadusastmeid ja *p*-väärtusi.

Within Subjects Effects

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p
Vihje täpsus	None	80215.524[a]	2.000[a]	40107.762[a]	69.206[a]	< .001[a]
	Greenhouse-Geisser	80215.524	1.339	59926.441	69.206	< .001
Residuals	None	44045.227	76.000	579.542		
	Greenhouse-Geisser	44045.227	50.866	865.915		

Note. Type III Sum of Squares
[a] Mauchly's test of sphericity indicates that the assumption of sphericity is violated (p < .05).

Saadud tulemuste raporteerimine käib analoogselt ühesuunalise ANOVA raporteerimisega, kuid korrigeeritud vabadusastmete raporteerimisel on kaks alternatiivi. Üks võimalus on raporteerida korrigeeritud vabadusastmed: F(1,3; 50,9) = 69,2; p < 0,001, η^2 = 0,646 või raporteerida korrigeerimata vabadusastmed koos Greenhouse-Geisseri väärtusega: F(2; 76) = 69,2; p < 0,001, η^2 = 0,646, GG=0.669 (vabadusastmed korrigeerimata). Viimase kasuks kallutab otsustama asjaolu, et korrigeeritud kujul vabadusastmed inimsilmale vähe informatiivsed. Kui korrigeerimata vabadusastmetega koos raporteerida ka Greenhouse-Geisseri väärtus, siis on igal ühel võimalik korrigeeritud vabadusastmed välja arvutada, sest nende saamiseks tuleb korrigeerimata vabadusastmed lihtsalt Greenhouse-Geisseri väärtusega läbi korrutada.

5. ÜLESANDED ISESEISVAKS HARJUTAMISEKS

Avage andmefail **maad.csv**.

Tegemist on riikide tasemel arvutatud indeksitega. Muutuja „SKP_grupid“ jagab riigid kolme rühma: madal, keskmine, kõrge. Tehke antud andmestikuga läbi järgmised ülesanded:

Ülesanne 1. Võrrelge innovatsiooniindeksit SKP gruppide lõikes. *Millist keskmiste võrdlemise testi kasutate? Mis on tulemused, st kas on erinevusi või mitte? Tehke ka post-hoc testid ja joonis.*

Ülesanne 2. Võrrelge demokraatiaindeksit SKP gruppide lõikes. *Millist keskmiste võrdlemise testi kasutate? Mis on tulemused, st kas on erinevusi või mitte? Tehke ka post-hoc testid ja joonis.*

Kasutage andmefaili **drugtrial.sav**.

Tegemist on ravimiuuringu andmestikuga, mis sisaldab muutujatena uuritava identifitseerimisindeksit, sugu, annust ning ravijärgset skoori. Andmestiku allikas: http://staff.bath.ac.uk/pssiw/stats2/page16/page16.html

Ülesanne 3. Uurige, kas esineb soo ja/või ravimi annuse peaefekt ja/või nendevaheline interaktsioon.

Kasutage andmefaili **recall_data.csv**.

Selles katses esitati inimestele erineva emotsionaalse väärtusega sõnu ja paluti neid hiljem meenutada. Näeme andmestikus, et iga katseisik on läbinud kõik katsetingimused („Pos“ – positiivse tähendusega sõnad, „Neg“ – negatiivse tähendusega sõnad, „Neu“ – neturaalse tähendusega sõnad).

Ülesanne 4. Uurige, kas meenutamise edukus sõltus katsetingimusest.

Statistiline modelleerimine
II loeng

Standardviga, usalduspiirid, statistiline olulisus, keskmiste võrdlemine

Karin Täht

23.09.19

Eesmärgiks loengus on anda ülevaade statistilise olulisuse mõistest, keskväärtuse usaldusspiiridest ning rühmade keskmiste võrdlemise mudelist.

- Meenutame varasemast tuttavaid mõisteid nagu valim ja üldkogum
- Vastame küsimustele mida tähendab järeldamine statistikas ja mis on standardviga
- Saame selgeks, kuidas teha kindlaks, et 2 või enama rühma keskmised mingi tunnuse lõikes erinevad üksteisest statistiliselt oluliselt
- Samuti vaatame, kuidas veenduda, et tuvastatud statistiliselt oluline erinevus on ka sisuliselt oluline (piisavalt suur).

Sissejuhatus

Üldkogum ja valim

- Üldkogum (populatsioon) on teatud nähtuste (objektide, isikute) hulk, mida soovitakse objektiivsete meetoditega tundma õppida (näiteks üliõpilased Eestis)
- Valimiks nimetatakse teatud hulka üldkogumi elemente, mille mõõtmisandmed on uurija käsutuses

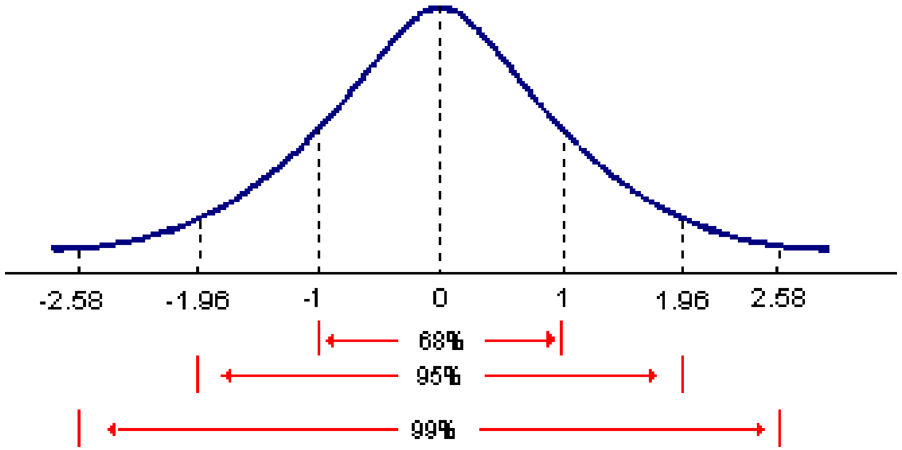
Kirjeldav statistika tegeleb valimi resümeerimisega, järeldava statistika ülesanne on üldistuste tegemine üldkogumi kohta

Meenutame eelmist loengut

Muutujate mõõteskaalad

	Näitab erinevust	Näitab erinevuse suunda	Näitab erinevuse suurust	Absoluutne nullpunkt
Nominaalskaala	X			
Järjestusskaala	X	X		
Intervallskaala	X	X	X	
Suhteskaala	X	X	X	X

Normaaljaotus (95% vaatlustest jääb vahemikku [-1.96;1.96])



Normaaljaotuslikkuse vaatamiseks kasulikud jaotuse omadused

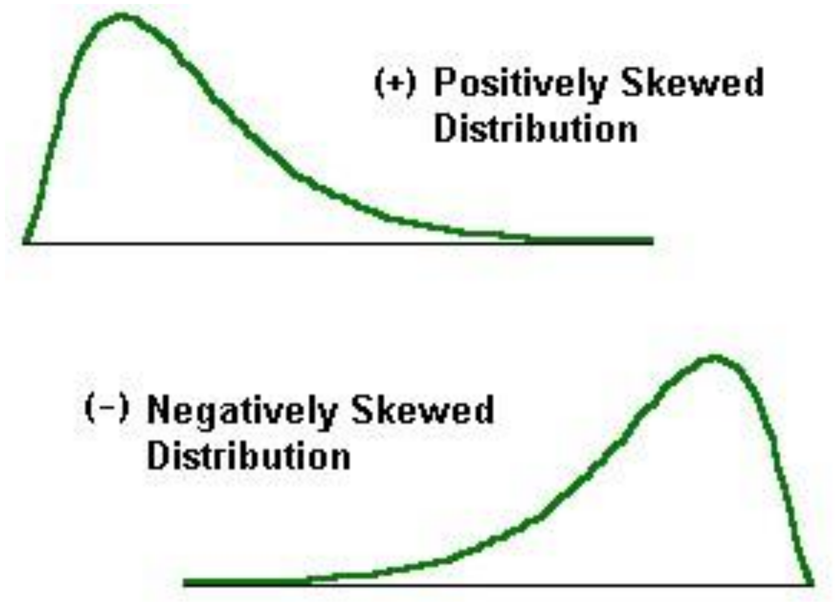
- asümmeetriakordaja (*skewness*)
  - kui (eba/)sümmeetriline on „kellukakõver“

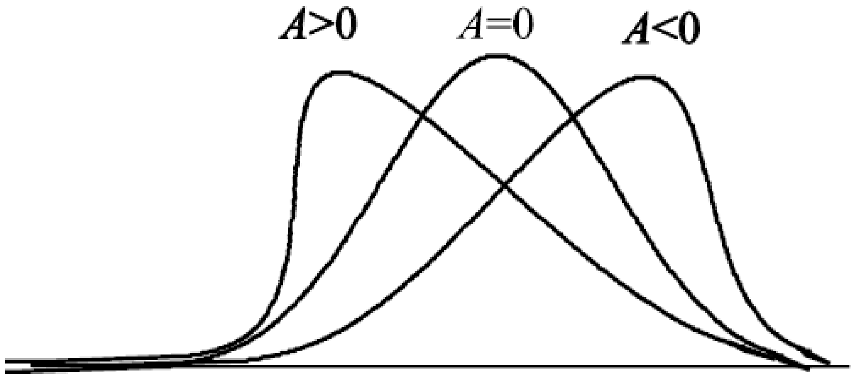
$$a = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i - m)^3}{s^3}$$

Kus **m** - aritmeetiline keskmine, $\mathbf{x_i}$ on tunnuse väärtused, **n** - valimi suurus, **s** - standardhälve

Asümmeetriakordaja

- Kokkuleppeliselt on tegemist normaaljaotusega, kui asümmeetriakordaja väärtus on vahemikus [-2; 2], konservatiivsemalt ka [-1; 1]





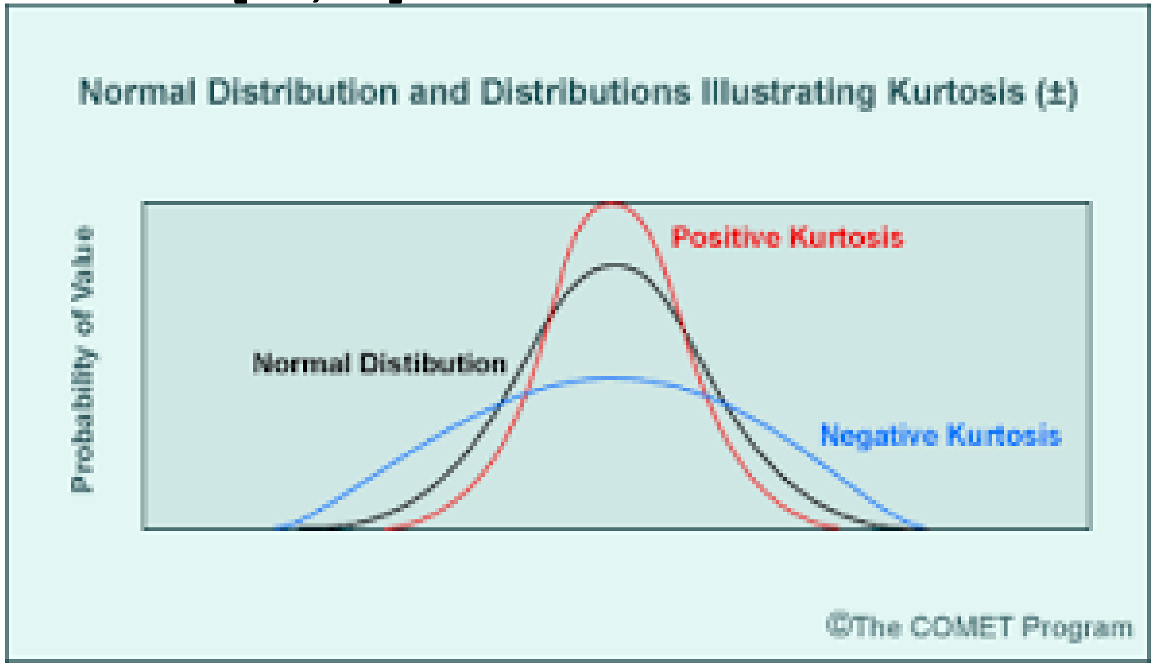
Normaaljaotuslikkuse vaatamiseks kasulikud jaotuse omadused

- järsakusaste, järsakusastmekordaja (*kurtosis*)
  - sagedusjaotuse tipu „teravus“, järsakus

$$e = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i - m)^4}{s^4} - 3$$

Järsakusastmekordaja (*kurtosis*)

- Kokkuleppeliselt on tegemist normaaljaotusega, kui järsakusastmekordaja väärtus on vahemikus [-2; 2], konservatiivsemalt ka [-1; 1]



Parameetriline ja mitteparameetriline statistika

Kui tunnus on arvuline ja ligilähedaselt normaaljaotusele, saame rakendada parameetrilist statistikat.

Järjestustunnustele, normaaljaotusest erinevate jaotuste puhul, väikeste valimite puhul on kohasem rakendada mitteparameetrilist statistikat.

See tuleneb asjaolust et erinevaid andmeanalüüsi meetodeid välja töötades on tehtud erinevad eeldused.

https://www.thoughtco.com/parametric-and-nonparametric-methods-3126411

Standardviga ja usalduspiirid

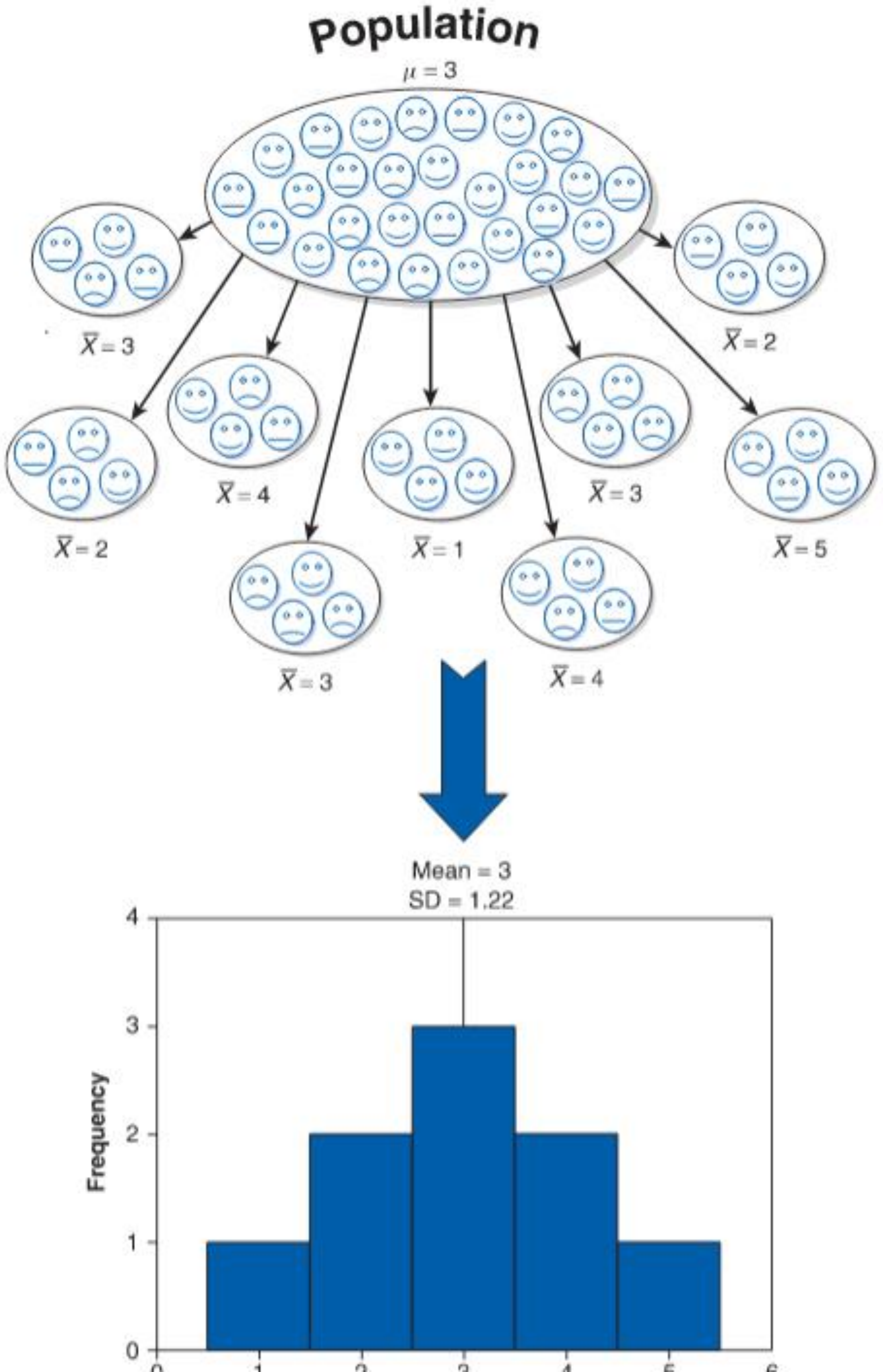
Üldkogum ja valimid
Valimite aritmeetilised keskmised

Tavaliselt tegutseb uurija ühe konkreetse valimiga.

Teoreetiliselt on üldkogumist võimalik “välja võtta” suur hulk **etteantud suurusega erinevaid** juhuslikke valimeid.

Iga sellise valimi jaoks saame leida tema **aritmeetilise keskmise**.

Tekib uus muutuja: ühest populatsioonist võetud valimite aritmeetilised keskmised. Viimastest saame moodustada omakorda tunnuse, millel on jaotus.



Population
$\mu = 3$
$\overline{X} = 3$
$\overline{X} = 2$
$\overline{X} = 4$
$\overline{X} = 2$
$\overline{X} = 1$
$\overline{X} = 3$
$\overline{X} = 5$
$\overline{X} = 3$
$\overline{X} = 4$
Mean = 3
SD = 1.22
Frequency

Olemasolevatest andmetest järelduste tegemine populatsioonile

Standardhälve näitab kui hästi keskmine esindab mõõdetud andmeid.

Enamasti pole eesmärk mitte valimi kohta võimalikult palju teada saada, vaid valimi põhjal teha järeldusi populatsioonile.

Siinkohal on oluline teada, kui suur **standard VIGA** (*standard error*).

Standard viga

Olgu meil k valimit ühest üldkogumist.

Nende valimite aritmeetilisi keskmised: m_1, m_2, …, m_k.

Võime tekkinud keskmisi vaadelda uue muutujana, millele on omakorda võimalik leida aritmeetiline keskmine ja standardhälve.

$$M = \frac{1}{k}\sum_{i=1}^{k} m_i$$

$$SE^2 = \frac{1}{k}\sum_{i=1}^{k} (m_i - M)^2$$

M - valimite jaotuse aritmeetiline keskmine,

SE – valimite jaotuse standardhälve e. standardviga

Valimite aritmeetiliste keskmiste jaotuse standardhälvet nim. STANDARDVEAKS.

$$M = \frac{1}{k}\sum_{i=1}^{k} m_i \qquad SE^2 = \frac{1}{k}\sum_{i=1}^{k}(m_i - M)^2$$

M - valimite jaotuse aritmeetiline keskmine,
SE – valimite jaotuse standardhälve e. standardviga.

Näide akadeemilise testi andmetega

Sõnavara	Diagrammid	Andmed	Informeeritus	tekst	matemaatika	ruumiline	AT
9	5	2	8	3	6	10	43,00
17	7	7	4	4	6	10	55,00
14	16	7	13	5	2	3	60,00
14	10	4	12	3	9	11	63,00
12	11	4	16	4	9	10	66,00
15	17	15	15	6	12	2	82,00
19	18	12	17	4	13	10	93,00
8	5	3	3	3	6	1	29,00
11	6	3	7	2	0	11	40,00
14	10	3	6	1	9	5	48,00
15	9	2	10	5	4	4	49,00
15	11	7	9	3	8	7	60,00
11	6	7	9	3	9	12	57,00
15	13	10	10	2	10	4	64,00
12	13	4	15	2	9	9	64,00
10	9	9	19	2	7	13	69,00
16	10	9	15	4	8	13	75,00
18	17	15	15	3	18	13	99,00
19	18	9	19	1	17	13	96,00
19	17	13	15	8	15	15	102,00
14	7	3	8	0	6	5	43,00
9	9	9	9	4	3	3	46,00

Seitse erinevat juhuslikku väljavõtet AT tegijate hulgast

	Minimum	Maximum	Mean		Std. Deviation	Skewness	Kurtosis
	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic
AT 1	21,1	82,4	46,93	1,76	14,54	,23	-,37
AT 2	21,1	104,6	51,81	2,10	17,41	,72	,68
AT 3	24,1	103,2	53,32	2,10	18,29	,90	,45
AT 4	22,2	103,2	49,91	1,90	16,19	,74	,97
AT 5	23,8	104,6	52,87	1,91	16,32	,70	1,16
AT 6	17,0	92,1	51,34	1,90	16,15	,25	-,29
AT 7	23,8	81,0	48,98	1,85	15,04	,25	-,88

Konkreetse valimi puhul arvutatakse: valimi standardhälve jagatud ruutjuurega valimi suurusest.

$$SE = \frac{s}{\sqrt{n}}$$

- s- standardhälve, n- valimi suurus
- mida suurem hajuvus valimis, seda suurem standardviga.
- standardviga saab vähendada, suurendades valimi suurust.
- Mida väiksem standardviga, seda kindlamad võime olla, et valimi aritmeetiline keskmine on lähedane üldkogumi keskväärtusele.

Järeldamine statistikas

- Valimi põhjal arvutatud arvkarakteristikute põhjal saame ainult hinnangud üldkogumi parameetritele.
- Viimast nimetatakse statistiliseks järeldamiseks.
- Millest sõltub statistilise järeldamise täpsus?

  - Mida suurem on valim, seda täpsem on meie hinnang üldkogumile.
  - Seega suurendades valimit, saame üldkogumi parameetreid hinnata täpsemalt.
  - KUID: me ei saa mitte kunagi 100% kindlusega öelda, et üldkogumi arvkarakteristik on võrdne mingi konkreetse arvnäitajaga.

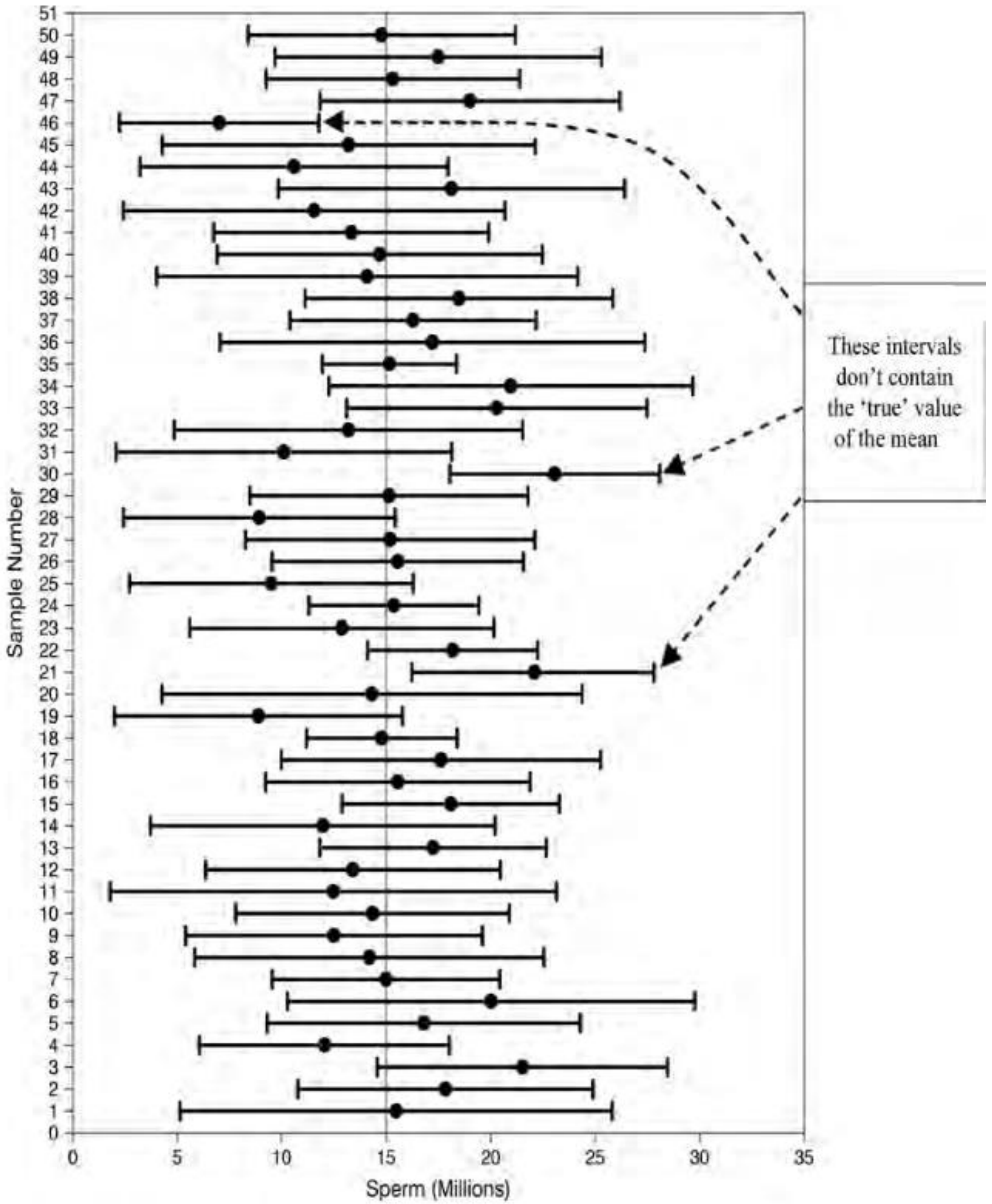
Keskväärtuse vahemikhinnang (usalduspiirid)

- Keskväärtus e üldkogumi aritmeetiline keskmine

- Üldkogumi keskväärtuse vahemikhinnanguks nim. valimi põhjal määratud vahemikku, kuhu valimi keskmine kuulub teatud (küllalt suure) tõenäosusega (enamasti 95%, aga ka 99%).

Järgnev joonis illustreerib ühest üldkogumist võetud 50 valimi keskmisi ja usalduspiire.

FIGURE 2.8 The confidence intervals of the sperm counts of Japanese quail (horizontal axis) for 50 different samples (vertical axis)



Usalduspiirid (*confidence interval*)(ka vahemikkhinnang)

Vastavalt usaldusnivoo väärtusele arvutatakse parameetri usalduspiirid so. Kaks arvu, mille vahel parameeter asub etteantud tõenäosusega.

Valem 95% usalduspiiride arvutamiseks:

Alumine usalduspiir= m-1.96*SE

Ülemine usalduspiir=m+1.96*SE

Olulisuse nivoo (*level of significance*)

- Öeldakse ka a -nivoo,
- See on tõenäosus, mille uurija ise määrab
- Öeldakse ka vea tõke, vea piir

Väga tihti võetakse olulisuse nivooks 0,05, see tähendab, et uurija annab endale aru, et kuna ta 100 protsendilise tõekindlusega ei saa midagi valimilt üldkogumile üldistada, siis ta „lubab endale eksida" 5 %.

Usaldusnivoo (*confidence level*)

95 % usaldusnivoo näitab, et 95 % tõenäosusega on tulemus usaldusäärne.

On oma olemuselt tõenäosus. Tähistatakse 1-α.

Olulisuse nivoo e. vea piiri tähistus: α.

Statistiline olulisus (*statistical significance*)- väljendab järelduste juhuslikkuse määra, teisisõnu tulemuste usaldusväärsust.

Keskmiste võrdlemine

Hüpoteeside testimine ehk keskmiste võrdlemine

Vajalik erinevate keskmiste erinevuse hüpoteeside testimisel:

- Kas ühe erakonna valijad on vanemad kui teise erakonna valijad?

- Kas naissoost üliõpilased saavad sõnavaratestis paremaid tulemusi kui meessoost üliõpilased?

....

Nendele küsimustele vastamiseks on alati põhimõtteliselt vastata järgmistele küsimustele:

kas rühmad (või valimid ja nende jaotused) on nii sarnased, et võime öelda, et nad kuuluvad samasse üldkogumisse? Või on nad nii erinevad, et esindavad kahte erinevat üldkogumit?

Hüpotees: naised tunnetavad ruumi temperatuuri jahedamana kui mehed


ScienceDirect Journals Books

Download PDF Export Search ScienceDirect Advanced search

Article outline Show full outline

Abstract
Keywords
1. Introduction
2. Materials and methods
3. Gender and thermal comfort
4. Gender and use of controls
5. Gender differences
6. Discussion
7. Conclusion
Acknowledgements
References

Figures and tables

Table 1
Table 2
Table 3
Table 4
Table 5
Table 6

ELSEVIER

Building and Environment

Volume 42, Issue 4, April 2007, Pages 1594–1603

Gender differences in thermal comfort and use of thermostats in everyday thermal environments

Sami Karjalainen

Show more

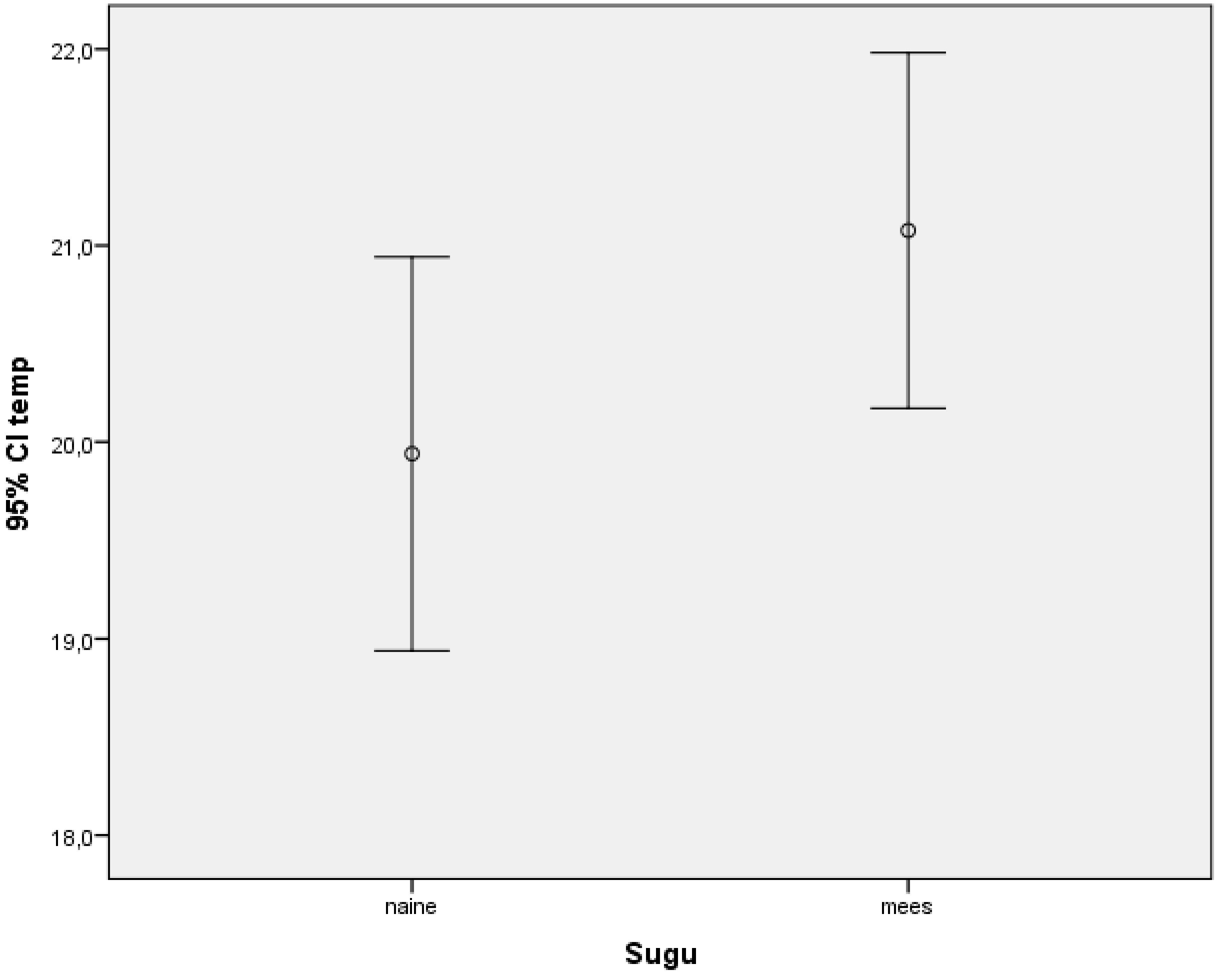
doi:10.1016/j.buildenv.2006.01.009

Abstract

Differences in thermal comfort between male and female subjects are generally considered to be small. In this study gender differences in thermal comfort and use of thermostats were examined by a quantitative interview survey with a total of 3094 respondents, and by controlled experiments. The studies were carried out in Finland and considered everyday thermal environments: homes, offices and a university. The results show significant gender differences in thermal comfort, temperature preference, and use of thermostats. Females are less satisfied with room temperatures than males, prefer higher room temperatures than males, and feel both uncomfortably cold and uncomfortably hot more often than males. Although females are more critical of their

http://www.sciencedirect.com.ezproxy.utlib.ee/science/article/pii/S0360132306000242

Näide: usalduspiirid meeste ja naiste pakutud keskmisele temperatuurile



	Mean	N	Std. Deviation
naine	19,94	25	2,43
mees	21,08	17	1,76
Total	20,40	42	2,23

Olulisuse testid e hüpoteeside kontrollimine

Kas erinevate rühmade keskmised on sedavõrd erinevad, et on põhjust pidada neid eri üldkogumitest pärinevaks.

Selleks tuleb läbida järgnevad sammud:

- püstitada nullhüpotees ning alternatiivne e. sisuline hüpotees
- defineerida testimise protseduur, sealhulgas olulisuse nivoo
- otsustada, millist keskmiste erinevuste testi kasutada
- arvutada teststatistikud ja nendega seotud olulisuse tõenäosused
- arvutada efekti suuruse näitajad
- teha järeldus, kas andmed on kooskõlas nullhüpoteesiga või mitte

Nullhüpotees ja alternatiivne hüpotees

- NULLHÜPOTEES: Alustatakse eeldusest, et valimid ei erine.

H_0: $\mu_1 = \mu_2$ (*nullhüpotees*)

- Kui erinevus kahe valimi aritmeetilise keskmise vahel osutub liiga suureks, et pidada neid valimeid ühest üldkogumist pärinevaks, lükatakse nullhüpotees ümber ja jäädakse ALTERNATIIVSE HÜPOTEESI juurde.

Olgu μ_1 esimese ja μ_2 teise üldkogumi keskmised

H1: $\mu_1 \neq \mu_2$ (kahepoolne alternatiivne hüpotees)

- Või: ühe alamkogumi keskväärtus on suurem kui teise oma

H1: $\mu_1 > \mu_2$ (ühepoolne alternatiivne hüpotees)

Hüpoteeside testimine kui statistiline mudel

Mis on meie mudeliks kui võrdleme kahe valimi keskmisi?

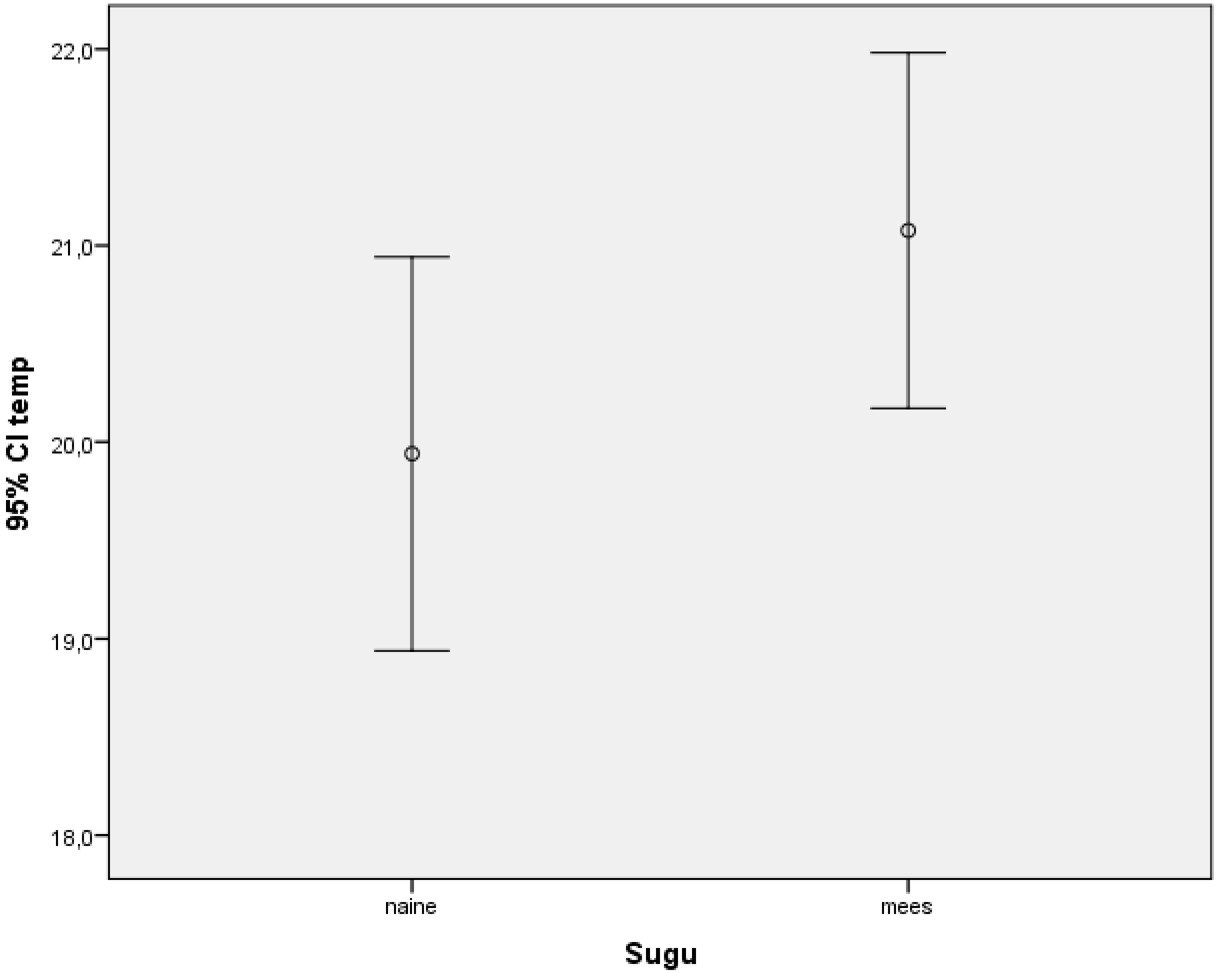
Olgu meil kaks valimit, ühe aritmeetiline keskmine X_1 ja teisel X_2

Standardhälbed vastavalt s_1 ja s_2.

Hüpoteeside testimise puhul oletame, et

valimite keskmiste erinevus on suurem kui oleks oodatud erinevus populatsioonide keskmiste vahel (mis on 0 kui nullhüpotees kehtib).

Näide 1: Mehed, naised ja temperatuur



	Mean	N	Std. Deviation
naine	19,94	25	2,43
mees	21,08	17	1,76
Total	20,40	42	2,23

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper
temp	Equal variances assumed	1,170	,286	-1,655	40	,106	-1,1365	,6865	-2,5239	,2510
	Equal variances not assumed			-1,759	39,778	,086	-1,1365	,6462	-2,4427	,1698

Kui suur on erinevus?

Efekti suurus näitab, kui suur on mõõdetud erinevus.

Efekti suuruse näitajaid on erinevaid (Cohen’i d, r) sõltuvalt millist keskmiste võrdlemise testi läbi viiakse. Nad on standardiseeritud suurused, seega saab ühes võrdlusülesandes välja arvutatud efekti suurust võrrelda teises võrdlusülesandes välja arvutatud (sama) efekti suurusega.

Efekti suurus

APA (*American Psychological Association*) range soovitus on efekti suuruse näitaja kõigile publikatsioonidele lisada.

Seega- leidnud statistiliselt olulise mõju (erinevuse), tuleb arvutada ka, kui suur see mõju (erinevus) on.

Suur või väike efekti suurus (Cohen'i d)?

Cohen'i d:

d < 0,2 – väike efekt

0,2 < d < 0,8 – keskmine ehk arvestatav efekt

d > 0,8 – suur efekt

R: 0- pole üldse efekti

1- täiuslik efekt

r ei ole lineaarsel skaalal (st 0,6 ei ole 2*0,3)

- r = .10 **(small effect):** In this case the effect explains 1% of the total variance.
- r = .30 **(medium effect):** The effect accounts for 9% of the total variance.
- r = .50 **(large effect):** The effect accounts for 25% of the variance.

Arvutame efekti suurus näites 1

http://www.uccs.edu/~lbecker/

Selleks on vaja rühmade keskmisi ja standardhälbeid:

	Mean	N	Std. Deviation
naine	19,94	25	2,43
mees	21,08	17	1,76
Total	20,40	42	2,23

D=0,54- tegu on keskmise efektiga.

Sisuliselt on tegu keskmiste erinevusega standardhälbe ühikus. Ehk keskmised erinevad üksteisest poole standardhälbe võrra.

Võrdlusülesannete liigitus 2 rühma korral

Oluline, kas võrreldavad rühmad on üksteisest sõltumatud või mitte

- Kui erinevad rühmad tekivad näiteks sootunnuste (või mingi rühmitava tunnuse) alusel, siis on rühmad üksteisest sõltumatud ja kasutatakse **t-testi sõltumatute rühmade** jaoks (*independent groups*)
- Kui võrreldakse ühe ja sama katsegrupi seisundeid erinevail ajahetkedel (kordusmõõtmised), siis on tegemist omavahel sõltuvate rühmadega ja kasutatakse **sõltuvate rühmade t-testi** (*dependent groups*)

Eeldused t-testi rakendamiseks

1.Andmed on vähemalt intervall skaalal

1.1 Sõltuvate rühmade t-testi puhul peab olema normaaljaotusega jaotus, mis on moodustatud järgmiselt: ühel korral mõõdetud tulemustest on lahutatud teisel korral mõõdetud tulemused.

1.2 Sõltumatu t-testi puhul peab tunnus rühmade lõikes olema ligikaudu normaaljaotusega, samuti peavad rühmade dispersioonide olema homogeensed ning valimid peavad olema sõltumatud

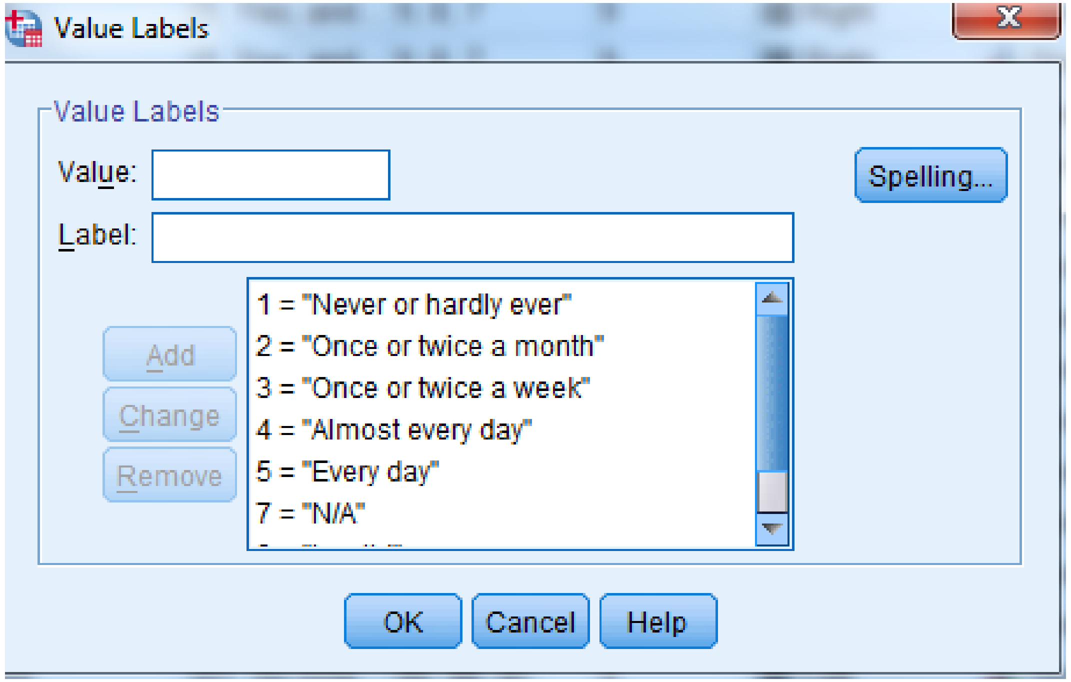
Näide 2: t-testi mitteparameetriline analoog

PISA 2012

Eesti valim

Kui tihti kasutad sotsiaalvõrgustikke?

Kas tüdrukud kasutavad tihedamini sotsiaalvõrgustikke kui poisid?



Keskmiste võrdlus kui parameetriliste testide eeldused pole täidetud

	Parametric	Non-parametric
Assumed distribution	Normal	Any
Assumed variance	Homogeneous	Any
Typical data	Ratio or Interval	Ordinal or Nominal
Data set relationships	Independent	Any
Usual central measure	Mean	Median
Benefits	Can draw more conclusions	Simplicity; Less affected by outliers
Tests		
Choosing	Choosing parametric test	Choosing a non-parametric test
Correlation test	Pearson	Spearman
Independent measures, 2 groups	Independent-measures t-test	Mann-Whitney test
Independent measures, >2 groups	One-way, independent-measures ANOVA	Kruskal-Wallis test
Repeated measures, 2 conditions	Matched-pair t-test	Wilcoxon test
Repeated measures, >2 conditions	One-way, repeated measures ANOVA	Friedman's test

http://changingminds.org/explanations/research/analysis/parametric_non-parametric.htm

Mitme rühmade keskmiste võrdlemine

Vaatame näiteks järgmise artikli teist tabelit:

Kwon, M., Kim, D. J., Cho, H., & Yang, S. (2013). The smartphone addiction scale: development and validation of a short version for adolescents. *PloS one*, *8*(12), e83558.

Table 2. Socio-demographic characteristics and SAS-SV scores.

Variables		N (%) Mean±SD	SAS-SV Mean±SD	p (N = 540)
Age		14.46±0.54		
Sex	Boy	343(63.5)	23.75±10.26	<.001
	Girl	197(36.5)	27.89±11.17	
Socio-economic state	Low	86(24.0)	25.08±10.65	.828
	Medium	191(53.2)	25.90±11.10	
	High	82(22.8)	25.94±11.12	
Family	Parents	442(81.9)	25.53±11.03	.173
	Single-parent	69(12.8)	23.07±9.62	
	No-Parent	23(4.3)	26.70±9.61	
Alcohol	Yes	21(4.1)	27.52±9.98	.295
	No	496(95.9)	25.03±10.69	
Smoking	Yes	22(4.2)	27.94±10.91	.210
	No	496(95.8)	25.03±10.61	
Time of smartphone use	Weekday	3.03±3.09		
	Weekend	4.07±4.08		
Purpose of smartphone use	Messenger and SNS	300(55.6)	26.29±11.07	.077
	Entertainment	114(21.1)	23.42±10.03	
	Web surfing	43(8.0)	24.79±9.88	
	Etc.	83(15.4)	24.29±10.86	
Self evaluation of smartphone addiction	Non-addiction	312(57.8)	21.82±8.34[a]	<.001
	Addiction	134(24.8)	34.02±10.34[ab]	
	Don't know	94(17.4)	24.17±11.68[b]	

Paarisvõrdlus *versus* mitmene võrdlus

Miks üldse mitmest võrdlust on vaja?

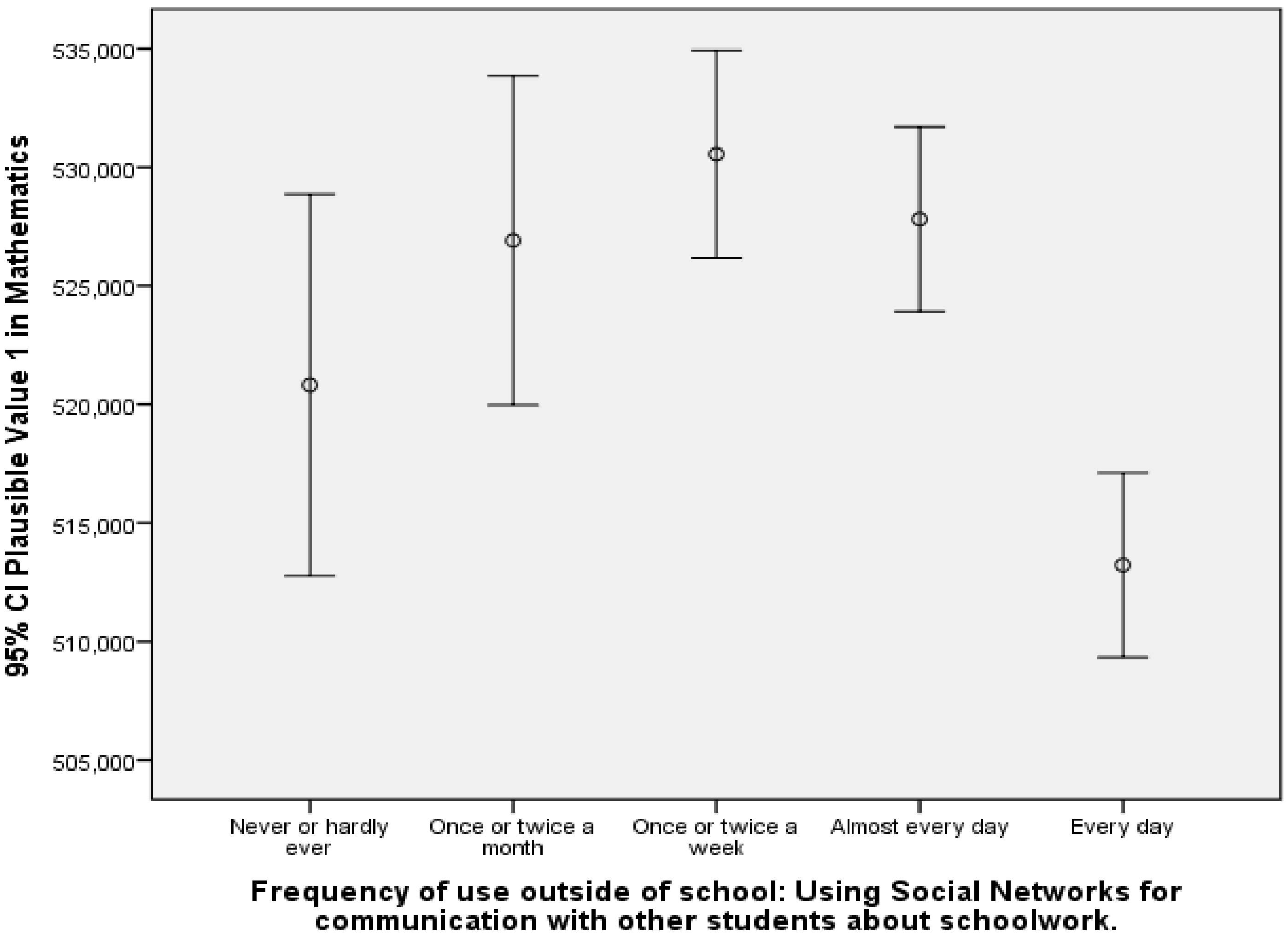
Viime mitu korda läbi paarisvõrdlusi!

Vaatame näidet: valim 15- aastased eesti õpilased, uuring PISA 2015

Soovime läbi viia 5 grupi võrdlust: 1. õpilased kes ei kasuta sotsiaalmeediat, 2. kasutavad 1 või 2 korda kuus, 3. kasutavad 1 või 2 korda nädalas, 4. kasutavad peaaegu iga päev, 5. kasutavad iga päev.

Uurimisküsimus: millisel rühmal on matemaatikatesti tulemus kõige madalam?

Näide: Matemaatikatesti tulemuste erinevused sõltuvalt sotsiaalmeedia kasutuse sagedusest (PISA 2015 andmed):



Mitmese testimise probleem

- Teame, kuidas läbi viia paarisvõrdlusi, oletame, et meil on rohkem rühmi, kas testime erinevusi neis siis paarikaupa?
- Olgu olulisuse nivoo 0.05 ja olgu meil 3 rühma (nende vastavad keskväärtuse hinnangud oleksid μ_1, μ_2, μ_3)
- Siis peaksime testima järgmisi erinevusi
- μ_1, μ_2
- μ_2, μ_3
- μ_1, μ_3

Kolm võrdlust

- Eelduse kohaselt igal sõltumatul võrdlusel on I tüüpi vea mittetegemise tõenäosus 0,95
- Kuna need võrdlused on sõltumatud, siis kokku kolme võrdluse I tüüpi vea mittetegemise tõenäosuse saamiseks, tuleb kolme omad korrutada
- 0,95*0,95*0,95=0,857, teisisõnu I tüüpi vea tegemise tõenäosus:

1-0,857=0,143 ehk 14,3 protsenti (seega oluliselt suurem kui 5%)

- Üldisemalt, kui meil on k rühma, siis on vaja teha

võrdlust

$$C = \frac{k!}{2(k-2)!}$$

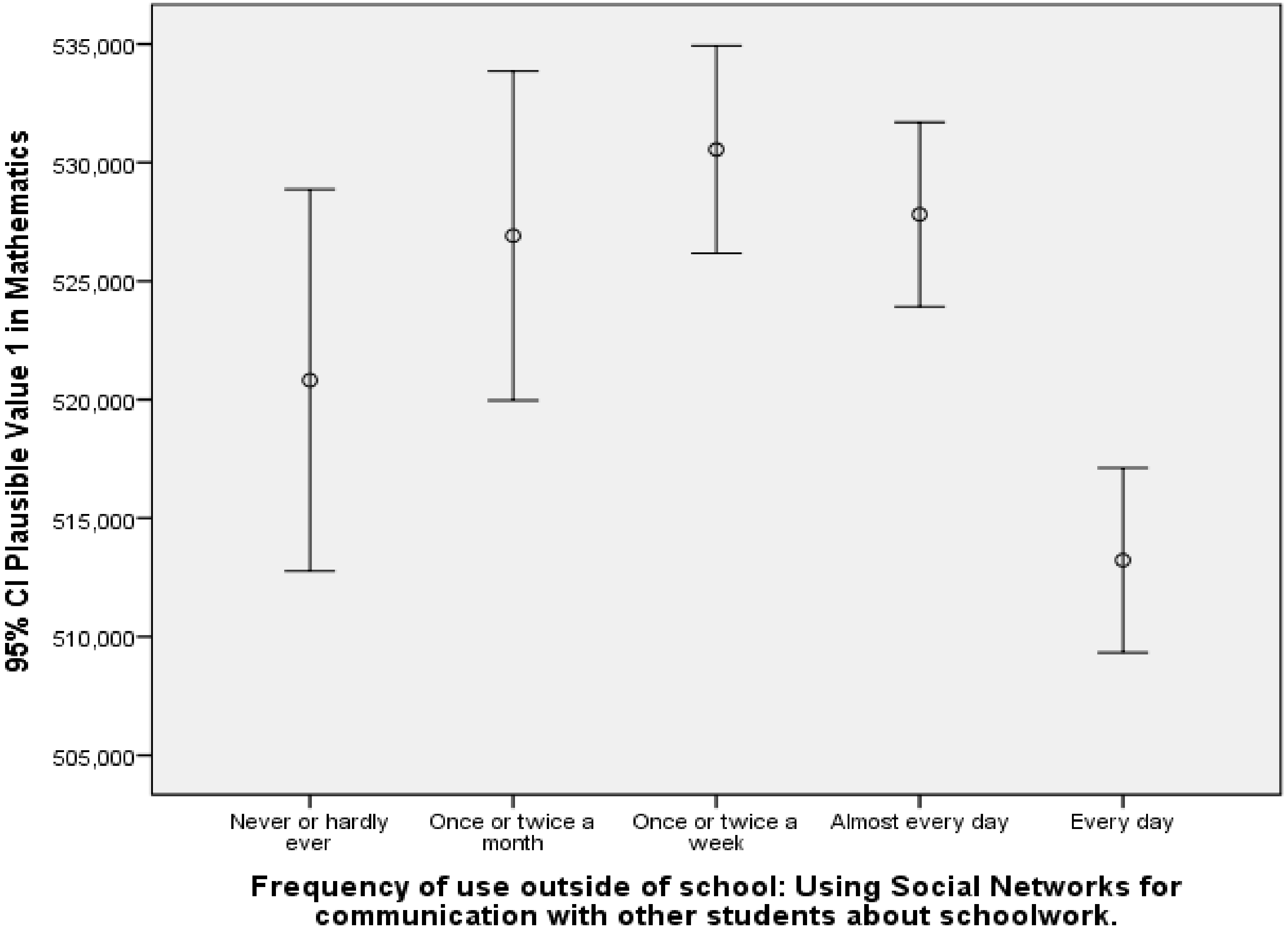
Ühesuunaline dispersioonanalüüs (*One-Way ANOVA*)

Erinevate rühmade keskmiste võrdlusülesanded on väga iseloomulikud sotsiaalteaduslikule andmeanalüüsile.

Mitmene võrdlus (enam kui kahe sõltumatu rühma võrdlemine):

1. kas üldse leidub erinevusi rühmade keskmiste vahel
2. konkreetselt milliste rühmade keskmiste vahel on statistiliselt erinevusi

Meenutades näidet: esmaselt püüame teada saada, kas üldse esineb erinevusi rühmade vahel?



Hüpoteesid

NULLHÜPOTEES. H_0: $\mu_1 = \mu_2 = \ldots = \mu_n$, ehk alamkogumite keskväärtused ei erine

SISUKAS HÜPOTEES, H1: leidub vähemalt üks alamkogumite paar, mille korral $\mu_i \neq \mu_j$

Nullhüpoteesi kehtimise tõenäosuse hindamine:

Eeldame, et nullhüpotees kehtib.

Siis $\mu_1 = \mu_2 = \ldots = \mu_n$

Ehk kõik grupid on ühest ja samast üldkogumist ning valimite hajuvuse põhjal saaks hinnata üldkogumi hajuvust (dispersiooni).

Dispersioonanalüüs (ANOVA)

- Tegeleb keskväärtuste erinevuse hindamisega (kuigi nime on saanud hajuvuse näitaja järgi, kuna keskväärtusi puudutava nullhüpoteesi kehtimist hinnatakse valimite hajuvuse kaudu).

- Võimalikud uurimisküsimused:

Kas katse tulemused on eri rühmades erinevad

Kas mingi manipulatsioon mõjutab testi tulemusi? Millisel tasemel?

Kas kool on testi tulemusi mõjutav faktor?

Kategoriaalset tunnust (sugu, vanus, kool) mille alusel gruppi jaotatakse tunnused nim. FAKTORIKS või FAKOTORTUNNUSEKS.

Sõltumatute rühmade dispersioonanalüüsi (ANOVA) eeldused

ANOVA on parameetriline test, tema läbi viimiseks peavad olema täidetud eeldused:

- Sõltuv tunnus peab olema vähemalt intervallskaalal
- Võrreldavad grupid peavad olema omavahel sõltumatud
- Tunnuse hajuvused võrreldavates gruppides peavad olema sarnased
- Tunnuse jaotus (gruppides eraldi võetuna) peaks olema ligilähedane normaaljaotusele.

Näide 3: olulised suurused dispersioonanalüüsi puhul

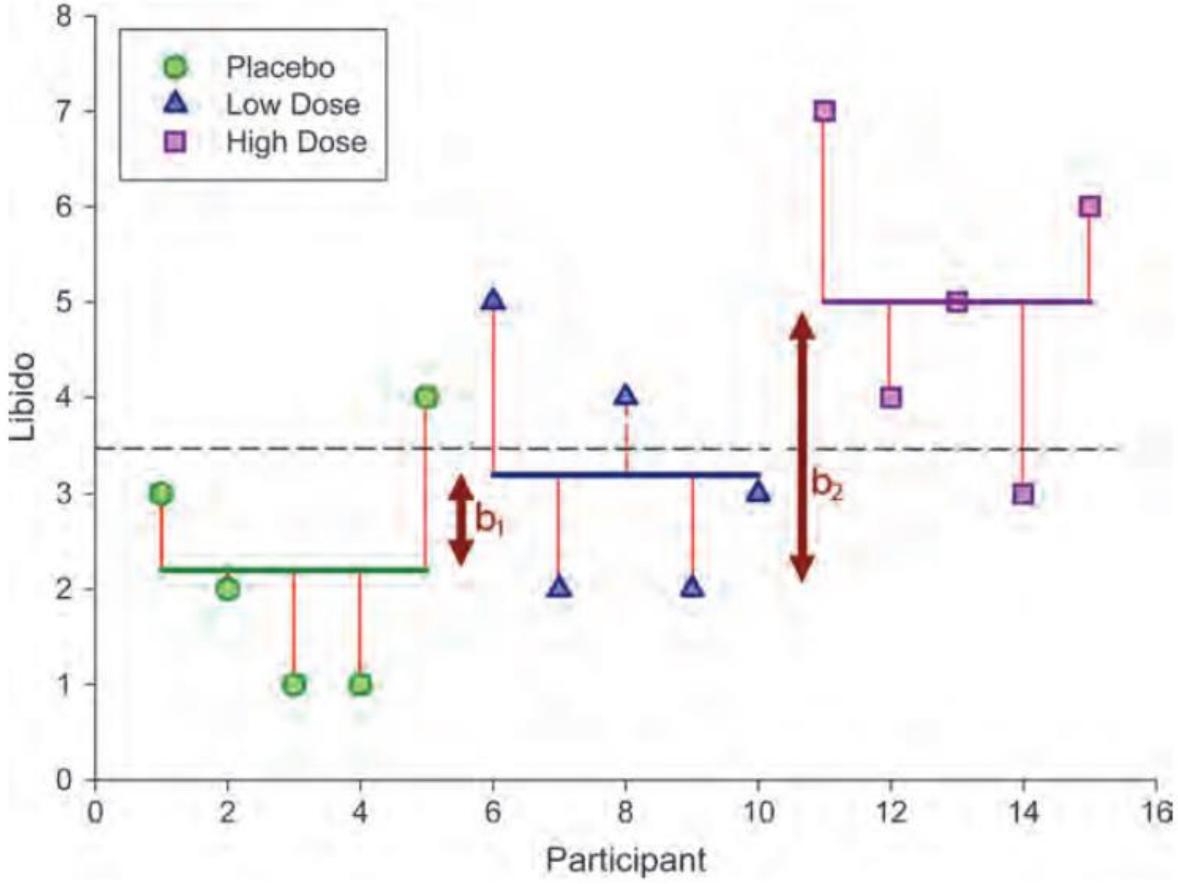
TABLE 10.1 Data in **Viagra.sav**

	Placebo	*Low Dose*	*High Dose*
	3	5	7
	2	2	4
	1	4	5
	1	2	3
	4	3	6
$\bar{X}$	**2.20**	**3.20**	**5.00**
s	**1.30**	**1.30**	**1.58**
s^2	**1.70**	**1.70**	**2.50**
	Grand Mean = **3.467**	Grand SD = **1.767**	Grand Variance = **3.124**

Rühmade keskmised, standardhälbed ja dispersioonid, samuti üldine keskmine (*grand mean*) standardhälve (*grand SD*) ja dispersioon (*grand dispersioon*).

Samad andmed graafiliselt kujutatuna

FIGURE 10.2 The Viagra data in graphical form. The coloured horizontal lines represent the mean libido of each group. The shapes represent the libido of individual participants (different shapes indicate different experimental groups). The dashed horizontal line is the average libido of all participants



Loogika, mis on mitme rühma keskmiste võrdlemise taga

- Kõige lihtsam mudel on- me saame andmeid iseloomustada üldise keskmisega (*grand mean*) (tähendaks, et efekti ei ole)
- Me võime andmetele sobitada mingi teise mudeli, kui see mudel sobitub andmetele hästi, peab see olema parem kui ainult üldise keskmise kasutamine
- Oletame, et selleks mudeliks on rühmade keskmised. Mida suurem on erinevus rühmade keskmiste vahel, seda suurem on erinevus mudeli ja keskmise vahel
- Kui erinevus rühmade vahel on küllalt suur- on see mudel parem kui andmete sobitamine üldise keskmisega
- Sellisel juhul on meie mudel (skooride ennustamine rühma keskmiste abil) parem, kui meil poleks mudelit (skooride ennustamine üldise keskmise järgi). Teisisõnu rühmade keskmised on oluliselt erinevad.

Erinevad *post-hoc* testid võrdsete dispersioonidega rühmade jaoks

Post hoc (lad. k) tähendab „peale seda“

Kasutatakse, et teha kindlaks, milliste rühmade keskmiste vahel täpsemalt erinevusi esineb, kui F suhe on näidanud, et üldine erinevus rühmade keskmiste vahel on olemas.

LSD (*least significant difference*)- nagu tehtaks eraldi vajalikud t-testid

Bonferroni- väga range test

Tukey’i test- nii see kui eelmainitu kontrollivad väga hästi I tüüpi vea tegemise tõenäosust, aga neil pole piisavalt statistilist jõudu.

Dunn ja Sheffe- umbes samaväärsed nagu kaks eelmist

REGWQ (Ryan, Einot, Gabriel and Welsch Q protseduur)- Fieldi õpiku järgi parim (kui rühmade suurused on võrdsed)

Erinevad *post-hoc* testid MITTE võrdsete dispersioonidega rühmade jaoks

Erinevates AT programmides võivad olla mõnevõrra erinevad,

SPSS pakub

Tamhane'i test ja Dunnett'i T3 ja C test, on kõik üsnagi konservatiivsed.

Games-Howell on kõige võimsam ja töötab ka kui rühmad on erineva suurusega. Võib olla liiga liberaalne väikeste valimite korral.

Efekti suuruse arvutamine ühesuunalise ANOVA korral:

One thing you will notice is that SPSS doesn't routinely provide an effect size for one-way independent ANOVA. However, we saw in equation (7.4) that:

$$R^2 = \frac{SS_M}{SS_T}$$

Of course we know these values from the SPSS output. So we can simply calculate r^2 using the between-group effect (SS_M) and the total amount of variance in the data (SS_T) – although for some bizarre reason it's usually called **eta squared, η^2**. It is then a simple matter to take the square root of this value to give us the effect size r:

Kokkuvõtes:

	Parametric	Non-parametric
Assumed distribution	Normal	Any
Assumed variance	Homogeneous	Any
Typical data	Ratio or Interval	Ordinal or Nominal
Data set relationships	Independent	Any
Usual central measure	Mean	Median
Benefits	Can draw more conclusions	Simplicity; Less affected by outliers
Tests		
Choosing	Choosing parametric test	Choosing a non-parametric test
Correlation test	Pearson	Spearman
Independent measures, 2 groups	Independent-measures t-test	Mann-Whitney test
Independent measures, >2 groups	One-way, independent-measures ANOVA	Kruskal-Wallis test
Repeated measures, 2 conditions	Matched-pair t-test	Wilcoxon test
Repeated measures, >2 conditions	One-way, repeated measures ANOVA	Friedman's test

Kuidas kasutada t-testi:
TUDENGI SPIKKER

1. Kas sõltumatul muutujal on **2 taset**?

Jah

Ei
→T-test pole Sinu jaoks!

2. Kas grupid on sõltumatud?

Jah → <u>INDEPENDENT SAMPLES T-TEST</u>

3. Kas sõltuv muutuja on vähemalt intervallskaalal?

Jah

Ei

Mann-Whitney U test

4. Kas tunnus on rühmade lõikes normaaljaotuslik?

Jah

Ei

Mann-Whitney U test

5. Kas rühmade hajuvused on homogeensed?

Jah

Ei

Studenti t-test

Welchi t-test

Ei → <u>PAIRED SAMPLES T-TEST</u>

3. Kas sõltuv muutuja on vähemalt intervallskaalal?

Jah

Ei

Wilcoxoni signed-rank test

4. Kas tunnuste keskmiste vahe on normaaljaotuslik?

Jah

Ei

Studenti t-test

Wilcoxoni signed-rank test