

TÖLVUTÍMAR Í AÐFERÐAFRÆÐI II

5. Verkefni í SPSS: Kassarit og *t*-próf

Í þessu verkefni munum við vinna með ný gögn sem innihalda eigindlega flokkabreytu og samfellda meginlega breytu. Fyrst skoðum við dreifingu samfelldu breytunnar út frá flokkabreytunni með því að fá bæði lýsandi mæligildi og myndir af dreifingunni.

Gögnin eru úr TIMSS (*Third International Mathematics and Science Study*), alþjóðlegri rannsókn á námsárangri í stærðfræði og náttúrufræði sem Íslendingar tóku þátt í árið 1999. Við munum vinna með upplýsingar um kynferði og frammistöðu í stærðfræði í litlu útaki íslenskra 9 ára barna.

Skoðun gagna með lýsandi mælitölum

Fyrsta skref í allri úrvinnslu er að skoða gögnin og átta sig á helstu eiginleikum þeirra. Við skulum byrja á því að fá lýsandi mælitölur yfir báðar breytur og mynd af dreifingu stærðfræðieinkunnar eftir kynferði.

Fyrst skaltu sækja gögnin. Þau finnur þú á K-drifi í möppunni K:/gba/spssdata. Gagnaskráin heitir Timss2.

Gagnaskráin inniheldur tvær breytur. Breytan *itsex* er eigindleg breyta með upplýsingum um kynferði nemandans og *aimatscr* er samfelld breyta sem sýnir frammistöðu nemandans á prófinu. Frammistaða barnanna er gefinn upp á stöðluðum alþjóðlegum mælikvarða.

Þar sem kynferði (*itsex*) tekur venjulega aðeins tvö gildi fáum við skýrar upplýsingar um dreifingu breytunnar með því að skoða tíðni (*Frequencies*) gilda hennar.

Skrifaðu **FRE Itsex** í skipanaglugga eða veldu *Frequencies* í valglugga.

Skoðaðu tíðnitöfluna vel. Oft sjást einhver óvenjuleg gildi sem taka þarf afstöðu til. Því er slík tíðnitafla oft mikilvægt fyrsta skref í úrvinnslu gagna. Það er mikilvægt að sjá öll sérkennilegheit í upphafi áður en eiginleg úrvinnsla hefst svo ekki þurfi að endurtaka úrvinnsluna.

Breytan *itsex* tekur tvö gildi. Gildið „1“ stendur fyrir stúlkur og gildið „2“ fyrir pilt. Öll önnur gildi eru brottfallsgildi.

Skoðaðu tíðnitöfluna og athugaðu hvort að einhver óvenjuleg gildi sé þar að finna. Skilgreindu síðan breytuna. Ósennileg gildi skalt þú flokka sem brottfall (*missing value*).

Tíðnitafla gefur ekki góða mynda af *Aimatscr* (stærðfræðieinkunn) þar sem hún er samfelld meginleg breyta. Því fengjum við í raun aðeins upptalningu á öllu úrtakinu. Hentugra er að skoða lýsandi mælitölur (*Descriptives*) fyrir breytuna.

Skrifaðu **DES Aimatscr** í skipanaglugga eða veldu *Descriptives* í valglugga.

Virðast einhver ósennileg gildi vera í *Aimatscr* þegar þú skoðar hæsta og lægsta gildi breytunnar? Ef þú værir að skoða þín eigin gögn myndir þú geta lesið í hæsta og lægsta gildið, séð hvernig þau eru tilkomin og tekið rökstudda afstöðu til þeirra. Í þessu tilfelli hefurðu tiltölulega litlar upplýsingar. Láttu þér nægja að skoða hvort

Kóðun á brottfalli

Hyggilegt er að kóða brottfall sem ósennilegt gildi miðað við önnur gildi breytunnar. Auk þess er gott að gefa því gildi sem er langt fyrir ofan eða neðan önnur gildi eða jafnvel sem neikvæða tölu þegar breytan tekur aðeins jákvæð gildi.

Slík gildi skera sig úr öðrum. Auk þess geta þau haft það mikil áhrif á niðurstöður að augljóst sé að eitthvað sé að. Það er betra en að þau hafi áhrif á niðurstöður án þess að eftir sé tekið.

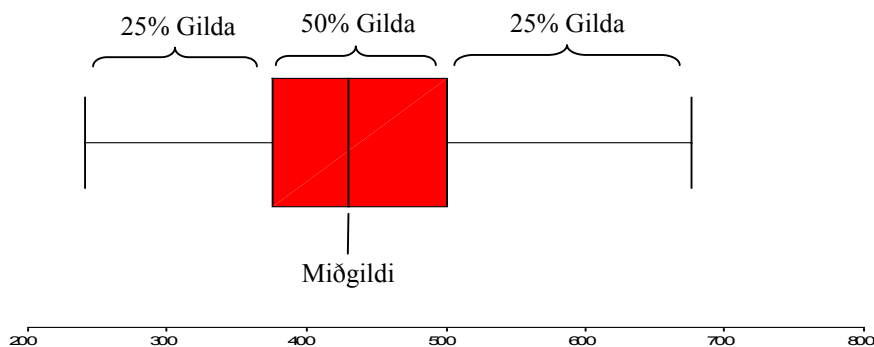
hæsta og lægsta gildið líti eðlilega út. Ef þau eru undarleg, er hugsanlegt að það þurfi að skilgreina brottfallsgildi. Hafðu samband við tölvukennara ef þú ert óviss.

Ljúktu við skilgreiningu breytunnar. Gefðu henni lýsingu og skilgreindu brottfallsgildi ef þörf krefur.

Ef það voru brottfallsgildi í gögnunum, ættirðu að biðja aftur um lýsandi mælitölur fyrir breytunnar tvær eftir að hafa skilgreint brottfallsgildin. Ef niðurstöðurnar virðast trúverðugar er ekkert því til fyrirstöðu að halda úrvinnslunni áfram.

Skoðun gagna með kassariti

Til þess að átta okkur frekar á sambandi kynferðis og stærðfræðieinkunnar skulum við skoða dreifingu stærðfræðieinkunnar eftir kynferði. Til þess skulum við fá kassarit (*boxplot*) af dreifingunni. Kassarit gefur mynd af dreifingu breytu út kringum miðgildi (*median*) hennar. Rifjaðu upp miðgildi (miðtölu; *median*) úr Aðferðafræði I.



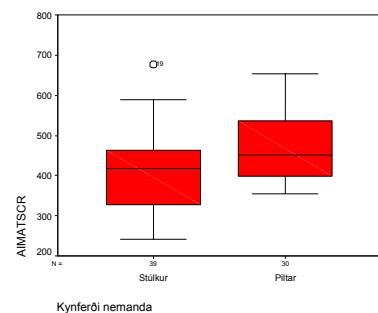
Myndin sýnir dæmigert kassarit. Mest áberandi er kassi með línu nálægt miðju kassans. Línan sýnir miðgildi dreifingarinnar. Hana má túlka svipað og meðaltal en er traustari (*robust*) en meðaltal gagnvart frávillingum (útlögum; *outliers*) og skekkju (*skewness*) í dreifingu en meðaltalið. Línan segir okkur því hvar miðju dreifingarinnar sé að finna.

Kassinn nær yfir helming mæligildanna. Ef kassinn er stuttur, nær þessi helmingur yfir lítið svæði; ef hann er langur, nær helmingur gagnanna yfir stórt svæði. Stærð kassans er því mælikvarði á dreifingu breytunnar. Við höfum notað dreifitölur (dreifni; *variance*) og staðalfrávik (*standard deviation*) til að meta dreifingu breyta; stærð kassans er annar mælikvarði.

Fyrir ofan kassann liggja 25% mæligildanna og 25% fyrir neðan hann. Út frá kassanum liggja strik, skeggin (*whiskers*), hvort í sína áttina. Stundum eru sýnd mæligildi með punktum eða opnum hringjum sem liggja fyrir utan skeggin. Þetta eru mæligildi sem ástæða er til að skoða nánar og gætu verið frávillingar (útlagar). Við höfum stundum kallað þetta jaðargildi; mæligildi sem eru yst á jöðrum breytunnar og hugsanlegir frávillingar.

Við skulum biðja um kassarit sem sýnir dreifingu breytunnar. Við viljum sjá hvort það sé munur eftir kyni og biðjum því um aðskilda kassa eftir kynjum.

Veldu *Analyse / Descriptive Statistics / Explore*. Færðu *Stærðfræðieinkunn (Aimatscr)* yfir á *Dependent List* og *Kynferði (Itsex)* yfir á *Factor List*. Farðu svo inni í *Plots* neðst á valmyndinni og taktu hakið af *Stem-and-leaf*. Smelltu á *Continue* og *OK* í aðalvalmynd.



Myndin sem þú færð á að vera áþekk myndritinu hér til hliðar. Þar má sjá hvernig dreifing stærðfræðieinkunnar er svipuð (svipuð stærð af kössum) hjá drengjum og stúlkum. Miðgildið er aðeins hærra hjá piltum en stúlkum og frammistaða þeirra því lítilliga betri. Eitt jaðargildi er hjá stúlkum með niðurstöðutöluna 677.

T-próf í tveimur hópum

Við viljum vita hvort frammistaða í stærðfæði hjá 9 ára börnum sé ólík eftir kynjum. Ef við hefðum aðeins áhuga á börnunum í úrtakinu, væri svarið fengið með myndinni. Við viljum hins vegar vita hvort frammistaða kynjanna sé ólík þegar litið er til allra 9 ára barna.

Spurningunni getum við svarað með t-prófi, en það getum við notað þegar við berum saman meðaltal tveggja hópa á samfelldri (eða nær samfelldri) meginlegri fylgibreytu. Prófið byggist á því að sett sé fram núlltilgáta (null hypothesis) um að enginn munur sé á meðaltölum hópanna. Niðurstaða prófsins segir síðan hvort munur meðaltalanna í úrtakinu sé sennilegur ef þessi núlltilgáta væri rétt.

Við skulum framkvæma t-próf til að athuga hvort munur sé á frammistöðu kynjanna í stærðfræði hjá íslenskum 9 ára börnum.

Veldu Analyse / Compare Means / Independent-Samples T Test. Færðu Stærðfræðieinkunn (AIMATSCR) undir Test Variable(s) og Kynferði (ltsex) undir Grouping Variable. Smelltu á Define Groups og skrifaðu 1 við Group 1 og 2 við Group 2. Smelltu á Continue og OK.

	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	Upper	
AIMATSCR	Equal variances assumed	,079	,779	-2,499	67	,015	-55,6576	22,27287	-100,114	-11,20076
	Equal variances not assumed			-2,539	65,621	,013	-55,6576	21,91820	-99,42340	-11,89178

Þú færð töflu svipaðri þeirri sem hér sést. Taflan er tiltölulega flókin og því mikilvægt að staldra aðeins við hana. Aðalhluti töflunnar hefur tvær línur. Í fyrsta dálki stendur „Equal variances assumed“ við efri línuna en „Equal variances not assumed“ við þá neðri. Þetta vísar til þess að þegar við berum saman meðaltöl í tveimur óháðum hópum með t -prófi, skiptir máli hvort staðalfrávik hópanna séu þau sömu eða ekki í þýði. Ef við álitum staðalfrávikin vera þau sömu, notum við efri línuna í töflunni en annars þá neðri.

Dálkar 2 og 3 gefa upp próf Levenes en það prófar tölfræðilega hvort gera megi ráð fyrir að staðalfrávikin séu eins eða ekki. Þriðji dálkur (merktur „Sig.“) gefur upp marktækt fyrir próf Levenes. Ef prófið er marktækt ($\alpha = 0,05$), álitum við að staðalfrávik hópanna séu ólík og notum neðri línu töflunnar. Ef prófið er hins vegar ómarktækt álitum við að staðalfrávik hópanna séu þau sömu og notum efri línuna.

Skodaðu p -gildið í þriðja dálk töflunnar.

Er próf Levenes marktækt? _____

Hvort ættirðu að nota efri eða neðri línu töflunnar? _____

Er hægt að hafna núlltilgátunni miðað við tvíhliða tilgátuprófun? _____

Svörin er að finna aftast í verkefninu.

Þessu næst þurfum við að lesa niðurstöðu t -prófsins út úr töflunni og athuga marktækt þess. Fjórði dálkur töflunnar sýnir niðurstöðu t -prófsins og sá fjórði sýnir frígráður þess. Við fáum tvær tölur og veljum efri eða neðri tölurnar eftir því hver niðurstaðan var í prófi Levenes. Efri tölurnar (efri línan í töflunni) á við ef próf Levenes var ómarktækt en þær neðri ef prófið var marktækt.

Skrifaðu niður niðurstöðu prófsins. Gættu þess að velja réttar tölur og gefa bæði upp prófniðurstöðuna (t -ið) og frígráðurnar.

Niðurstaða tölfræðiprófsins var: _____ (svör aftast).

Ef niðurstaðan er marktæk getum við hafnað núlltilgátunni ($T_0: \mu_{piltar} = \mu_{stúlkur}$) og tekið upp aðaltilgátuna ($T_0: \mu_{piltar} \neq \mu_{stúlkur}$). Það er misjafnt hvort við setjum tilgáturnar fram formlega eða ekki; í prófum í Aðferðafræði II er að sjálfsgöðu áskilið að formlegar og nákvæmar tilgátur séu settar fram.

Ef niðurstaða prófsins er ómarktækt getum ekki hafnað núlltilgátunni. Þá getum við í sjálfu sér lítið sagt; núlltilgátu má hafna en hún verður aldrei samþykkt.

Segðu nú til hvort niðurstaða prófsins sé marktæk eða ekki og hvaða tölfræðilega ályktun þú getur dregið um aðal- og núlltilgátuna. Miðaðu við tvíhliða próf.

Tölfræðileg túlkun á niðurstöðu t -prófsins: _____

Auk tölfræðilegrar túlkunar er mikilvægt að geta gert sér grein fyrir niðurstöðunni í almennu máli. Þetta er það sem kallað er efnisleg túlkun. Efnisleg túlkun er einfaldlega það að koma orðum að niðurstöðunni svo allir skilji. Mundu að efnisleg túlkun snýst ætíð um þýðið.

Settu fram efnislega túlkun á niðurstöðunum.

Efnisleg túlkun niðurstaðnanna: _____

Nú höfum við fengið endanlega niðurstöðu miðað við þær tilgátur sem tölfræðiprófið byggðist á. Það þarf auðvitað ekki að þýða að niðurstaðan sé *rétt*; við vitum ekki um kynjamun í þýði, heldur ályktum við um hann. Niðurstaðan er hins vegar *réttlætianlega*, þ.e. fengin með viðurkenndum aðferða- og tölfræðilegum aðferðum.

Þar sem við höfum engar aðrar upplýsingar, hljótum við að telja að niðurstaðan sé bæði *rétt* og *réttlætianleg*. Við álitum því að 9 ára stúlkur standi sig ver í stærðfræði en piltar. Þá er eðlilegt að spyrja sig hversu miklu ver þær standi sig. Það eru ýmsar kenningar um hvað valdi misjafnri frammistöðu kynjanna; líffræði-, tilfinninga- og félagsálfræðilegar skýringar hafa verið settar fram. Til dæmis er bent á að stúlkur fá litla hvatningu í stærðfræði meðan þyki mikilvæg grein fyrir pílta.

Lykilatriði er að vita hvort munurinn sé umtalsverður eða ekki. Við vitum að það munar töluverðu á kynjunum í úrtakinu en hversu mikill munur er á stærðfræðikunnáttu kynjanna miðað við öll íslensk 9 ára skólabörn?

Við búum svo vel að fá öryggisbil frá SPSS. Öryggisbilið er í tveimur öftustu dálkum töflunnar. Hér þarftu aftur að velja línu miðað við marktæktina á prófi Levenes.

Skoðaðu öryggisbilið og túlkaðu niðurstöður þess.

Túlkun á öryggisbilinu: _____

Öryggisbilið getur haft áhrif á efnislegu túlkunina. Þótt prófið sé marktækt og því kynjamunur hjá íslenskum 9 ára börnum, þarf sá munur ekki að vera umtalsverður. Ef

hann er lítill, skiptir hann auðvitað ekki neinu máli. Ef hann er hins vegar mikill, er það bæði áhugavert en einnig mikið áhyggjuefni. Þetta er það sem nefnt er áhrifastærð (*effect size*), þ.e. hve mikill munur (áhrif) eru í þýði.

Settu fram nýja efnislega túlkun þar sem þú tekur tillit til öryggisbilsins. Miðaðu við að þú sért að kynna niðurstöður fyrir almenningi eða einhverjum sem hugsanlega muni taka ákvarðanir byggðar á þinni túlkun.

Efnisleg túlkun þar sem minnst er á áhrifastærð: _____

Svör

Þar sem p -gildið fyrir próf Levenes er 0,779 er það mjög fjarri því að vera marktækt. Því gerum við ráð fyrir að staðalfrávik hópanna séu þau sömu í þýði og notum efri línu töflunnar („Equal variances assumed“).

Niðurstaða t -prófsins er gefin upp svona: $t(67) = -2,5$, $p < 0,05$. Þú velur efri tölurnar, þar sem próf Levenes var ómarktækt. Þú setur frígráðurnar innan svigans og prófniðurstöðuna á eftir samasemmerkinu. Yfirleitt er nóg að gefa upp einn aukastaf í niðurstöðunni. Algengast er að gefa marktæktarstigið (p -gildið) upp miðað við stöðluð viðmið (0,05; 0,01; 0,001). Ef einhver vill nákvæmari marktækt getur hann flett því upp sjálfur. [Ef próf Levenes hefði verið marktækt hefði niðurstaðan verið gefin upp svona: $t(65,6) = -2,5$, $p < 0,05$.]

Niðurstaða prófsins er marktæk miðað við $\alpha = 0,05$. Við getum hafnað núlltilgátunni og tekið upp aðaltilgátuna að það sé munur á stærðfræðikunnáttu kynjanna. Ef það væri enginn munur á kunnáttu kynjanna (í þýði) væri minna en 5% líkur á því að fá jafnmikinn eða meiri kynjamun í kunnáttu en þann sem er í úrtakinu.

Níu ára stúlkur á Íslandi standa sig ver í stærðfræði en piltar. [Þessa efnislegu túlkun má orða á ótal vegu.]

Öryggisbilið gefur til kynna að kynjamunurinn sé á bilinu -11 til -100 . Miðað við 95% öryggi gæti frammistaða stúlkna verið á bilinu $11 - 100$ stigum lægri en hjá piltum. Við getum borið þetta saman við staðalfráviknið en það er á bilinu $85-96$ eftir því hvort kynið er skoðað. Kynjamunur er mjög óverulegur ef miðað er við efri mörk bilsins $(-11,2)$ en mjög umtalsverður (rúmlega eitt staðalfrávik) ef miðað er við neðri mörk þess $(-100,1)$.

Miðað við niðurstöður okkar standa 9 ára stúlkur sig ver í stærðfræði en piltar. Við getum hins vegar lítið sagt um það hvort munurinn sé umtalsverður eða ekki. Þær upplýsingar sem við höfum benda til þess að það geti verið óverulegur og nánast enginn munur á frammistöðu kynjanna. Sömu upplýsingar gefa til kynna að munurinn gæti líka verið mjög mikill og skipt sköpum í öllu skólstarfi. Við getum ekki gert upp á milli þessara möguleika nema með frekari rannsóknnum og gagnasöfnum. [Þessa hugsun má orða á ótal vegu og hún þarf jafnvel ekki að vera svona ítarlega og nákvæmlega orðuð.]