Aðferðafræði II 10.05.03

Spurt og svarað

Hér finnur þú svör við fyrirspurnum í þriðja þriðjungi námskeiðsins sem nemendur hafa sent inn á árunum 1998–2003.

III. Ályktun í tveimur hópum, kíkvaðrat og fleira

Geturðu útskýrt öll skrefin í tilgátuprófun?

Getur þú farið í gegnum ferilinn á glæru 5 í Ályktunum í tveimur hópum, þ.e. atriði 1–6 og útskýrt það skref fyrir skref?

Þessi atriði hafa verið til umfjöllunar í gegnum allan annan þriðjung námskeiðsins. Eftirfarandi er yfirlit yfir hvernig skrefin sex tengjast öðrum hlutum námskeiðsins. Um leið tengja þau saman ólík efnisatriði fyrirlestra og glæra.

Skrefin sex er að finna á glæru 5 í Ályktunum í tveimur hópum. Önnur samantekt, en með færri skrefum, er á glæru 1 í Ályktunum í einum hópi (sjá einnig A: 154–155). Farið er í gegnum öll þrepin á glærum 2–5 í sama fyrirlestri. Inn á einstök atriði er síðan komið nokkrum sinnum á öðrum glærum.

Athugum lið fyrir lið hvernig þrepin tengjast öðrum hlutum námskeiðsins:

Skref 1: Setja fram aðal- og núlltilgátu.

Um þetta er rætt í Grunnatriði í ályktunartölfræði, glærur 14–17 með viðeigandi tilvísunum í kennslubók.

Hér er aðalatriðið að gera sér grein fyrir að stundum vill maður fá svör við tilteknum spurningum. Aðaltilgátuna er ekki hægt að prófa beint því hún tiltekur ekki nein ákveðin þýðisgildi; því þarf að setja fram núlltilgátu.

Skref 2: Velja marktektarstig.

Um þetta er fjallað í Ályktanir í einum hópi, glæra 6; A: 173-176. Við veljum marktektarstig miðað við þá áhættu sem við erum tilbúin að taka þegar núlltilgátan er rétt. Ef α= 0,05 og núlltilgátan er rétt, þá munum við (ranglega) hafna núlltilgátunni í 5% þeirra tilvika.

Skref 3: Framkvæma rannsóknina og fá upplýsingar úr úrtakinu

Rannsóknin hefur þann tilgang að fá úrtaksupplýsingar; á grunni þeirra reynum við að svara spurningunni sem felst í aðaltilgátunni.

Skef 4: Reikna *t*- eða *z*-próf

Þetta höfum við rætt oft, sjá m.a. Ályktanir í einum hópi, glæra 10. Sjá einnig Spurt og svarað: „Hvenær er *t*-dreifing notuð?

Skref 5: Bera saman vendigildi eða reikna p-gildi

Hér athugum við hvort úrtaksniðurstaðan sé ólíkleg undir núlltilgátunni. Ef hún er ólíkleg, sbr. skref 2, þá höfnum við núlltilgátunni og tökum upp aðaltilgátuna.

Skoðaðu Grunnatriði, glæra 17; A: 157-159. Sjá einnig Spurt og svarað: „Hvað þýðir p=0,05?“ og „Hvernig fæ ég p-gildi fyrir *t*- og kíkvaðratpróf?“ Um þetta er einnig fjallað um þetta á fleiri stöðum í glærum og fyrirlestrum.

Skref 6: Túlka niðurstöður

Hér túlkum við tölfræðilega og efnislega. Tölfræðileg túlkun snýst um það að hafna eða hafna ekki núlltilgátunni og því sem þeirri ákvörðun fylgir. Einnig þarf að segja nákvæmlega hvað felist í niðurstöðunni, þ.e. *p*-gildinu og höfnun (eða ekki höfnun) núlltilgátunnar.

Efnisleg túlkun felst í því að segja niðurstöðuna á mannamáli, svo amma skilji. Einnig getur verið gott að hugsa sér efnislegu túlkunina sem undirfyrirsögn í Morgunblaðinu.

Þið hafið fengið bæði efnislega og óefnislega túlkun í fyrirlestrum fyrir hver og eitt tölfræðipróf sem þar hefur verið framkvæmt. Einnig má styðjast við Spurt og svarað t.d.: „Marktekt og núlltilgátur,“ „Hvenær er kíkvaðrat- og *t*-próf marktækt,“ „Marktekt og núlltilgátur,“ „Tilgátuprófun og marktekt“ og „Hvað er efnisleg og tölfræðileg túlkun?“

Þetta verður að nægja í bili. Eins og þú sérð, krefst fullnægjandi svar við spurningunni þess að allt námsefnis sé tekið saman í svarinu. Skoðaðu þessar tilvísanir og sjáðu hvort þær hjálpa; það munu eflaust vakna sérhæfari spurningar eða spurningar um svörin. Ekki hika við að senda þær og ég mun leitast við að svara.

2000-03-17 GBA

Parað *t*-próf

Ég er í smá vandræðum með að skilja alveg formúlurnar á glæru 16 í sambandi við parað *t*-próf. T.d. í formúlunni fyrir staðalfrávik mismunar, hvaða tölur stendur *D* fyrir.

Önnur spurning kannski alveg út í hött en getur maður fundið út staðalfrávik mismunar ef maður hefur ekki tölurnar úr hópunum aðeins meðaltöl og staðalfrávik hópanna tveggja?

Það er fjallað um parað *t*-próf á bls. 226–229 í Agresti (kafli 7.4) og táknin útskýrð.

Táknið *D* stendur fyrir mismunatölu innan hvers pars (eða einstaklings) fyrir sig. Ef ég vildi t.d. athuga hvort munur væri á einkunnum í hlutaprófi I og II (auðvitað hef ég allt þýðið og þarf enga tölfræði, en horfum fram hjá því) þá myndi ég fyrir hvern nemanda reikna mismuninn á einkunn hans í prófi I og prófi II. Þessi tala væri táknuð með *D* og ég fengi eina tölu fyrir hvern nemanda. Athugaðu að sömu nemendur taka bæði prófin, þeir sem taka aðeins annað prófið detta út.

Þar sem ég er kominn með mismunatöluna *D* fyrir alla nemendur er ekkert í veginum fyrir því að reikna meðaltal og staðalfrávik fyrir *D*. Þetta er auðvitað aðeins hægt ef ég hef aðgang að niðurstöðum fyrir einstaka nemendur; ef ég hef aðeins hópmeðaltölin get ég ekki reiknað mismunatölurnar né staðalfrávik hennar.

Setja má tilgátur í pöruðu t-prófi fram á tvo ólíka vegu. Ég set tvíhliða tilgátur yfirleitt fram svona: T_0: $\mu_2 - \mu_1 = 0$. Agresti kýs að setja þær fram svona, sem hugsanlega er skýrara: T_0: $\mu_D = 0$.

1999-04-26 GBA

Hvernig fæ ég p-gildi fyrir t- og kíkvaðratpróf?

Hvernig breytir maður útkomu úr *t*-prófi í *p* gildi, svo að maður geti borið það við marktektar-mörkin?

Það eru í aðalatriðum tvær aðferðir við að kanna hvort niðurstaða prófs sé mark-tæk. Annars vegar get ég sett fram marktektarmörk (α) og flett upp vendigildi en hins vegar get ég fundið nákvæmalega líkindin undir núlltilgátunni (*p*-gildið). Ekki skiptir máli hvora aðferðina þú nota.

Fyrri aðferðin.

Setjum nú sem svo að við notum $\alpha = 0{,}05$ og fáum niðurstöðu úr *t*-próf sem er 2,93 miðað við 21 frígráðu. Til að kanna hvort prófið sé marktækt þyrfti ég að fletta upp vendigildinu í töflu B í Agresti. Miðað við tvíhliða próf, $\alpha = 0{,}05$ og 21 frígráðu gefur taflan upp vendigildið 2,080. Það þýðir að ef núlltilgátan er rétt eru 5% (α) líkur á að fá *t* sem er 2,080 eða hærra. Þar sem niðurstaðan 2,93 er hærra en vendigildið 2,080 er 5% eða minni líkur á niðurstöðunni undir núlltilgátunni og því er prófið marktækt og ég hafna núlltilgátunni.

Seinni aðferðin.

Ég get líka reiknað út líkindin á niðurstöðunni undir núlltilgátunni. Ef ég fram-kvæmi prófið í SPSS, þá prentar forritið út nákvæmt *p*-gildi. Ef ég er að reikna þetta í höndunum, þarf ég að nota töflu B.

Niðurstaðan var 2,93. Ég vel línuna sem á við 21 frígráðu í töflu B. Ég fer út línuna þar til ég finn hæsta vendigildið sem er jafnstórt eða lægra en niðurstaðan. Í þessu tilfelli er niðurstaðan hærri en öll vendigildin en hæsta vendigildið er 2,831. Nú skoða ég hvaða líkindi eru efst í viðkomandi dálki og finn táknið $t_{0,005}$. Þar sem ég miða við tvíhliða próf, þarf ég að tvöfalda líkindin svo þau verða 0,01. Ég veit því að rétta *p*-gildið er lægra en 1% og birti það svona: $p < 0{,}01$. Praktískt séð reyni ég að nota stöðluð *p*-gildi, þ.e. 5%, 1% eða 0,1%.

1999-04-15 GBA

Einhliða og tvíhliða próf

Hver er munurinn á einhliða og tvíhliða t-prófum. Af hverju er stundum notað „=“ en stundum „>“ í tilgátum?

Í tvíhliða próf er notuð stefnulaus (*nondirectional*) tilgáta. Dæmi um þetta er ef tilgátan er að karlar og konur hafi ólíka líkamshæð: T_1: $\mu_{KK} \neq \mu_{KVK}$. Þar sem andhverfan þessa er að þýðismeðaltölin sé þau sömu notum við samasemmerki í núlltilgátunni: T_0: $\mu_{KK} = \mu_{KVK}$. Núlltilgátan er því röng hvenær sem þýðismeðaltölin eru ekki nákvæmlega þau sömu.

Nú væri eðlilega að álíta að karlar séu að meðaltali hærri en konur. Því mætti setja fram þá stefnutilgátu að þýðismeðaltal karla sé hærra en þýðismeðaltal kvenna: T_1: $\mu_{KK} > \mu_{KVK}$. Andhverfa þessa er að karlar séu að meðaltali annað hvort jafnháir konum eða lægri en konur: T_0: $\mu_{KK} \leq \mu_{KVK}$.

Stefnulausa tilgátu prófum við með tvíhliða prófi. Gerum ráð fyrir því að líkamshæð kvenna og karla séu ólík í einhverju tilteknu úrtaki (úrtaksmeðaltölin ekki þau sömu). Þá gætum við prófað muninn með t.d. *t*-prófi; ef niðurstaðan er marktæk, myndum við hafna núlltilgátunni, hvort sem karlar væru hærri en konur eða konur hærri en karlar.

Stefnutilgátuna prófum við með einhliða prófi. Miðað við okkar tilgátur gætum við aðeins hafnað núlltilgátunni ef úrtaksmeðaltal karla er hærra en úrtaksmeðaltal kvenna. Ef konur reyndust hærri en karlar í úrtakinu, þá gætum við ekki hafnað núlltilgátunni—jafnvel þótt við fengjum marktækt *t*-próf.

1999-04-12a GBA

Aðskildar og samlagðar dreifitölur

Hver er munurinn á aðskildum og samræmdum dreifitölum. Við vitum að ef próf Levenes gefur *p* hærra en 0,05 að þá eigi að nota samræmdar og öfugt. En hvað þýðir þetta? Hvað erum við búin að sýna fram á ef að við samkvæmt niðurstöðum Levenes prófs ákveðum að hafa samræmdar dreifitölur eða öfugt.

Próf Levenes prófar þá núlltilgátu að enginn munur sé á dreifitölum hópanna. Ef það er marktækt höfnum við núlltilgátunni og gerum ráð fyrir að dreifitölurnar séu ólíkar.

Það hefur kosti að nota samræmdar (samlagðar) dreifitölur. Samlagðar dreifitölur gefa nákvæmari mynd af staðalvillu, þar sem við höfum tvær dreifitölur (staðalfrávik í öðru veldi) sem báðar eru óskekkt spá um sameiginlegt þýðisstaðalfrávik (og þar með staðalvillu einnig). Staðalvillan er ónákvæmari ef úrtaksstaðalfrávikin byggja á ólíkum þýðisstaðalfrávikum (aðskildar/aðgreindar dreifitölur); þessi ónákvæmni er ástæðan fyrir þeirri leiðréttingu frígráða sem notuð er fyrir aðskildar dreifitölur.

1999-04-12b GBA

Hvort nota ég aðgreindar eða samlagðar dreifitölur?

Þegar ég reikna t-próf í höndum og er ekki með niðurstöðu fyrir próf Levenes, hvernig veit ég hvort miða á við aðgreindar eða samlagðar dreifitölur?

Ég skil ekki af hverju Agresti gerir ráð fyrir aðgreindum dreifitölum í dæmi 7.7 (þegar staðalfrávik hópanna eru nokkuð svipuð) en samlögðum í dæmi 7.19 (þegar þau eru ekki eins). Er miðað við úrtaksstærð eða hvað?

Þótt vitað sé hvernig mismunur tveggja meðaltala dreifist annars vegar þegar þýðisstaðalfrávikin eru eins og hins vegar þegar þau eru ólík, eru deildar meiningar um hvenær skuli nota aðskildar og hvenær samlagðar dreifitölur.

Í Aðferðafræði II kynnum við tiltölulega vélgengt kerfi þar sem próf Levenes er notað til að skera úr um hvora aðferð skuli nota. Röklega séð er þetta ekki mjög fullnægjandi þar sem við gætum ýmist gert mistök af tegund I (talið þýðisstaðalfrávikin ólík þegar þau eru í reynd eins) eða mistök af tegund II (talið staðalfrávikin eins þegar þau eru í reynd ólík). Athuganir sýna þó að þessi leið virkar vel í reynd.

Önnur leið er að íhuga hvor möguleikinn sé sennilegri, þ.e. að þýðisstaðalfrávikin séu eins eða að þau séu ólík. Ein leið er að byggja á fræðilegum forsendum og meta þannig hvor möguleikinn sé líklegri. Í nær öllum tilfellum er þó eitthvert mið tekið af úrtaksstaðalfrávikunum. Ef þau eru mjög ólík, eykur það sennileika þess að þýðisstaðalfrávikin séu einnig ólík. Augljóslega er slíkt mat þó mjög huglægt.

Agresti tekur mið af því að ef úrtak er stórt verða niðurstöður svipaðar hvort sem notaðar eru samlagðar og aðgreindar dreifitölur. Því mælir hann með því að nota aðgreindar dreifitölur ef önnur hvor hópastærðin er 20 eða stærri. Þetta byggir á því að honum finnst auðveldara að nota þá formúlu heldur en formúluna fyrir aðgreindar dreifitölur. Ef hóparnir eru litlir mælir hann með því að byggja á samlögðum dreifitölum nema ef úrtaksstaðalfrávikin eru mjög ólík, t.d. annað þeirra tvöfalt hærra en hitt. Ef hóparnir eru litlir og staðalfrávikin mjög ólík mælir hann með aðgreindum dreifitölum og að frígráður séu leiðréttar með nálgun Welch-Satterthwaite.

Ef þú hefur ekki próf Levenes er skynsamlegast að nota aðgreindar dreifitölur ef staðalfrávikin eru mjög ólík. Leiðréttingu Welchs og Satterthwaites myndirðu aðeins nota ef hóparnir eru litlir. Að öðru leyti má fara aðrar leiðir en Agresti, t.d. máttu nota samlagðar dreifitölur í stórum úrtökum í stað þess að nota aðgreindar dreifitölur eins og Agresti kýs að gera. Í reynd skipta slíkar ákvarðanir litlu fyrir niðurstöðuna nema þegar unnið er með litla hópa.

Í dæmi 7.7 eru 152 einstaklingar í öðrum hópnum en 140 í hinum. Þetta eru því stórir hópar. Í slíkum tilfellum kýs Agresti að nota aðgreindar dreifitölur (en án leiðréttingar Welchs og Satterthwaites). Það skýrir hvaða formúla er gefin upp í svörunum. Í dæmi 7.19 eru 13 nemendur í öðrum hópnum en 17 í hinum. Þetta eru því litlir hópar samkvæmt viðmiðum Agrestis. Hann kýs að nota samlagðar dreifitölur í litlum hópum nema úrtaksstaðalfrávikin gefi annað til kynna. Annað staðalfrávikið er 2,1 og hitt er 3,2 og því er minna en helmingsmunur á staðalfrávikunum. Agresti er því sjálfum sér samkvæmur þegar hann kýs að nota samlagðar dreifitölur.

Finnst þér þetta ruglingslegt? Það er ekki furða. Aðalatriðið er þó þetta: Notaðu aðgreindar dreifitölur ef úrtaksstaðalfrávikin eru mjög ólík og sérstaklega ef hóparnir eru litlir. Ef þú hins vegar hefur niðurstöðu úr prófi Levenes, er einfalt og tiltölulega traust að láta hana ráða ákvörðuninni.

2001-04-23a GBA

Staðalfrávik í verkefni 84

Í verkefni 84 í **Stoðkveri** þarf maður að reikna staðalvillu fyrir *t*-próf, en hvar finnur maður upplýsingar um staðalfrávik fyrir formúluna s_1 og s_2?

Formúlu fyrir staðalfrávik finnur þú á formúlublaði fyrir annan hluta námskeiðsins.

Láttu það ekki rugla þig að oftast stendur s í formúlum fyrir staðalvillu. Það er engu að síður gert ráð fyrir að notað sé spágildi, þ.e. $\hat{\sigma}$. Mælitölurnar s_1 og s_2 standa þá fyrir spágildi staðalfrávika í hópum eitt og tvö, þ.e. $\hat{\sigma}_1$ og $\hat{\sigma}_2$.

Eftir að þú hefur reiknað spágildin fyrir hvorn hóp fyrir sig, setur þú þau í formúlu fyrir samlagt staðalfrávik, sbr. **Formúlur í þriðja þriðjungi**. Þetta sameiginlega staðalfrávik notarðu síðan í formúlunni fyrir staðalvillu fyrir mismun meðaltala.

Glæran **Samlagðar dreifitölur** í fyrirlestrinum **Ályktanir í tveimur hópum** sýnir hvernig þetta er reiknað.

2003-01-11j GBA

Hvenær er kíkvaðrat og *t*-próf marktæk?

Ef vendigildið er hærra en reiknað gildið fyrir t.d. kíkvaðrat eða *t*-próf, er niðurstaðan þá ómarktæk? Ef ekki hvernig er þá hægt að sjá að niðurstaðan sé ómarktæk.

Já. Þú flettir upp vendigildinu í töflu B eða C miðað við viðeigandi frígráður og alfamörk.

Gerum ráð fyrir að alfa sé 0,05. Ef núlltilgátan er rétt, þá verður niðurstaða tölfræðiprófsins (kíkvaðratprófsins eða *t*-prófsins) jafnhá eða hærri en vendigildið í 5% tilvika. Þegar niðurstöðurnar ná þessu marki, höfnum við núlltilgátunni og tölum um að prófið sé marktækt. Í þessum skilningi skilgreina þau alfamörk sem við veljum þá áhættu sem við erum tilbúin að taka þegar núlltilgátan er rétt.

1999-04-12c GBA

t dreifing og *t*-próf

Glósum mínum um marktekt í prófi Levenes ber ekki alveg saman og því væri gott að fá úr því skorið hvort ég hafi nú á endanum skilið þetta rétt.

Er það rétt skilið að þegar próf Levenes gefur í SPSS Sig.< 0,05 þá er það marktækur munur, staðalfrávikin ólík og við notum aðskildar dreifitölur?

Já, þetta er rétt athugað.

Í reynd erum við að prófa núlltilgátuna að dreifitölur hópanna séu eins í þýði. Ef próf Levenes er marktækt þá er munur dreifitalnanna í hópunum það mikill í úrtakinu að það er ólíklegt að fá þetta mikinn eða meiri mun ef dreifitölurnar eru eins í þýði. Því höfnum við því að dreifitölurnar séu eins í þýði og gerum ráð fyrir að þær séu ólíkar.

Ef dreifitölur eru ólíkar í þýði notum við t-próf fyrir aðskildar (aðgreindar; unequal; unpooled) dreifitölur, þ.e. línuna sem merkt er „equal variances not assumed“ í útprentununum úr SPSS.

1999-04-07 GBA

Hvað þýðir það „Prob > |T|“ 0,3071 í orðum í dæmi 7.22 í Agresti? Hvað er þessi tala að segja mér og hvar þetta í SPSS?

Samkvæmt töflu 7.11 er niðurstaða *t*-prófsins 1,1711 og líkurnar á þetta háu eða hærra *t* er u.þ.b. 0,31 ef núlltilgátan er rétt. Þetta samsvarar dálkinum sem er auðkenndur með „sig. (2-tailed)“ í útprentunum á *t*-prófi úr SPSS.

1999-04-07 GBA

Hvernig reikna ég *p* gildið í dæmi 7.28 í Agresti og hvað er *t* 2,67 (er það vendigildið)?

Talan 2,67 er niðurstaða *t*-prófsins. Þú flettir upp í töflu B með viðeigandi frígráðum og alfamörkum til að fá samsvarandi vendigildi. Í dæminu er ætlast til þess að þú gefir niðurstöðuna upp sem viðeigandi *p* gildi (t.d. $p < 0{,}05$) en það gerirðu með því að staðsetja rétta línu í töflunni miðað við viðeigandi frígráður og finna dálk þar sem niðurstaða *t*-prófsins er hærra en vendigildið en lægra en vendigildið í næsta dálki við.

1999-04-07GBA

Er útreikningur paraðs *t*-prófs alltaf einhliða?

Nei, parað *t*-próf getur verið ýmist einhliða (stefnutilgáta) eða tvíhliða (stefnulaus tilgáta).

1999-04-07 GBA

t-próf og samfelldar megindlegar breytur

Í verkefni 5 er talað um að við getum notað t-próf þegar við berum saman meðaltal tveggja hópa á SAMFELLDRI EÐA NÆR SAMFELLDRI MEGINDLEGRI BREYTU.

Þýðir þetta að það er ekki alltaf hægt að nota t-próf til að bera saman meðaltöl tveggja hópa? Fer það eftir því hvers konar breytu við erum með?

Tja, hvernig hefði átt að orða þetta. Við erum að koma á móts við óskir nemenda og fylgja langri hefð með því að tilgreina þær tegundir marktektaprófa og tengslastuðla sem henta fyrir hverja tegund breyta fyrir sig. En þú hefur hins vegar rétt fyrir þér þegar þú fullyrðir að þú getir notað *t*-próf hvenær sem það er merkingarbært að reikna meðaltal.

Meðaltöl eiga fyrst og fremst við þegar breytur eru samfelldar eða nær samfelldar. En það eru ótal undantekningar frá þessu og alvanalegt t.d. að reikna meðaltöl fyrir raðbreytur, t.d. fyrir svör við spurningum á spurningalistum (t.d. hægt að velja 1, 2, 3, 4 eða 5). Hér er aðalatriðið að vita hvað maður er að gera en ekki að fylgja einhverjum kreddum, hvorki frá mér né öðrum.

Þú ert sem sé búin að koma upp um meginveikleikann á öllum þessum töflum og leiðbeiningum sem við erum að gefa út. Mín vörn er af tvennu tagi: (1) Ég geri þetta hálf nauðugur (2) Svona leiðbeiningar geta hjálpað nemendum að taka byrjendaskrefin. Á móti kemur að slíkar leiðbeiningatöflur hamla líka, því það eru ekki allir færir um að sjá einfaldleikann að baki eins og þú hefur gert hér.

2000-04-29a GBA

Hvernig eru frígráður fengnar í dæmi 86?

Ég var að spá í dæmi 86 í Stoðkverinu. Þar erum við beðin um að reikna frígráður tólf para. Í dæminu var tekið fram að niðurstöður fyrir hvorn hópinn voru mældar, sem gefur til kynna að hér hafi verið að um tvo hópa ræða.

Er ekki rétt að frígráðurnar eru reiknaðar $n_1 + n_2 - 2$? Í svörunum hins vegar virðist eins og að frígráðurnar séu reiknaðar út frá einum hópi eða $n - 1$.

> Hér er um tvo háða hópa að ræða og því er notað parað *t*-próf. Nánar er fjallað um þetta efni í lok fyrirlestursins Ályktanir í tveimur hópum og í kafla 7.4 í Agresti.
>
> Í pöruðu *t*-prófi er það parið sem er stakið en ekki einstaklingurinn. Hvert par hefur sína niðurstöðutölu, þ.e. mismun í fjölda mótbára (þ.e. hversu oft ósammála). Í dæminu eru pörin 12 og hvert par með sína niðurstöðutölu. Tólf stök í eins hóps *t*-prófi gefur 11 frígráður; á sama hátt gefa tólf pör í pöruðu *t*-prófi einnig 11 frígráður.
>
> Láttu það ekki rugla þig að tala er um *hópa*. Parað *t*-próf er bæði notað þegar tvær mælingar, t.d. fyrir og eftir mælingar, eru gerðar á hverjum einstaklingi og þegar notaðir eru tveir hópar þar sem hver einstaklingur í öðrum hópnum er paraður við einstakling í hinum.

2002-04-23a GBA
2003-01-11h GBA

Túlkun *t*- og kíkvaðratprófs

Hér koma nokkur atriði sem okkur finnst vera óljós eftir að hafa farið yfir öll dæmin og lesið námsefnið. Ef hægt væri að koma þessu atriðum á hreint þá yrðum við þakklátir.

Túlkun fyrir *t*-próf.

Reiknað t er t.d. 2,34 sem er hærra en vendigildið 2,086 miðað við alfamörk 0,05 og 20 frígráður. Því höfnum við núlltilgátu, þ.e. munur er á þýðismeðaltölunum; t.d. áhorf fólks á sjónvarp fer eftir kyni í þýði.

Já, en auk þess eru minna en 5% líkur á þessum mun milli kynja eða meiri í úrtakinu ef núlltilgátan er rétt (ef engin kynjamunur er í þýði). Við höfnum núlltilgátunni og tökum upp aðaltilgátuna að munur sé á kynjunum í þýði.

2000-04-29b GBA

Hvernig yrði sambærileg túlkun fyrir kíkvaðratpróf?

Niðurstaða prófsins er hærri en vendigildið og því eru minna en 5% (þ.e. α, oftast 5%) líkur á þetta miklum eða meiri tengslum breytanna í úrtakinu ef engin tengsl eru milli þeirra í þýði (þ.e. ef núlltilgátan er rétt). Því hafna ég núlltilgátunni að engin tengsl (breyturnar séu óháðar) í þýði og tek upp aðaltilgátuna að ein breytan (t.d. kyn) hafi áhrif á hina (t.d. tíðni sjónvarpsáhorfs).

2000-04-29b GBA

Það sem vakir aðalega fyrir okkur er að okkur finnst ekki nógu skýrt hvað það er sem nauðsynlega þarf að koma fram í túlkun á þessum atriðum. Væri möguleiki á því að fá nákvæma útlistun á því hvað þarf að taka fram og hvað ekki svo túlkun sé fullnægjandi í svörum.

> Þetta og auk þess efnisleg túlkun. Ekki læra mitt orðalag utanað heldur skuluð þið reyna að átta ykkur á hvað nákvæmlega felist í því. Þetta eru grundvallaratriði við túlkun niðurstaða sem þurfa að koma fram. Minna máli skiptir nákvæmlega hvernig það er orðað svo fremi að orðalagið sé tæknilega rétt.

Vor 1999

Hvað er hermilíkan?

Hvað er hermilíkan og hvað segir það mér eða gerir fyrir mig??

> Hermilíkan er eftirlíking af þýði yfirleitt á formi tölvuforrits. Það gerir mér kleift að taka fjölmörg úrtök úr þekktu þýði og sjá þannig hvernig úrtakstölurnar hegða sér.
>
> Með hermilíkani er t.d. hægt að sjá að ef þýði er ekki normaldreift en ef úrtakið er nægjanlega stórt—þá munu úrtaksmeðaltölin engu að síður dreifast samkvæmt normaldreifingu. Þetta er markgildissetningin umtalaða; hana er hægt að staðfesta með hermilíkani. Þetta er því ein gagnsemi hermilíkana.

Vor 2000 GBA

Kíkvaðrat dreifing

Hvernig túlka ég töflur C í Agresti?

Til dæmis 10 frígráður og líkindin 0,050 gefur mér 18,31. Á þessi tala bara við halann hægra megin, eða þarf ég að hugsa um vinstri hlið kúrfunnar líka?

Við höfum aðeins áhuga á hægri hlið ferilsins. Vinstri hliðin samsvarar niðurstöðum í úrtaki sem eru líklegar ef núlltilgátan er rétt; hægri hlið ferilsins samsvarar niðurstöðum í úrtaki sem eru því ólíklegri eftir því sem við förum lengra til hægri á ferlinum.

Ef rauntíðni er jöfn væntitíðni verður kíkvaðrat 0,0. Eftir því sem rauntíðnin verður ólíkari væntitíðninni hækkar niðurstaða kíkvaðratprófsins. Ef frígráðurnar eru 10 og tilviljunarúrtak er dregið úr þýði með viðkomandi væntitíðni, þá eru aðeins 5% líkindi á því að fá það mikil frávik (mismun raun- og væntitíðni) í úrtakinu að kíkvaðratið verði 18,31 eða hærra.

1999-04-07a GBA

Hvernig er lesið úr kíkvaðrattöflu?

Hvernig les ég út úr kíkvaðrattöflu að eitthvað sé ólíklegt undir núlltilgátunni, til dæmis þegar kíkvaðrat er 9,49? Hvað er þetta að segja mér, hafna ég núlltilgátunni eða ekki?

Ef skilyrtu dreifingarnar eru mjög ólíkar jaðardreifingunni (þ.e. rauntíðnin er ólík væntitíðninni) þá verður kíkvaðrat hátt. Slík frávik verða alltaf einhver í úrtaki jafnvel þótt engin tengsl séu milli breytanna í þýði (þ.e. þó skilyrtu dreifingarnar séu eins og jaðardreifingin í þýði). Ef frávikin eru mjög mikil (kíkvaðratið mjög hátt), þá geta frávikin verið orðið það mikil að þau væru ólíkleg ef úrtakið væri úr þýði án tengsla.

Við segjum því að prófið sé marktækt (miðum oftast við 5% líkindi) og höfnum því að engin tengsl séu milli breytanna í þýði (höfnum núlltilgátunni).

Vor 2000 GBA

Frígráður kíkvaðrats

Ég er eitthvað að vandræðast með það hvernig eigi að finna frígráður fyrir kíkvaðrat. Ég er með formúluna en ég veit ekki hvernig á að finna *c* og *r* í formúlunni?

Bókstafurinn *c* stendur fyrir fjölda dálka (***c**olumns*) og *r* fyrir fjölda lína (***r**ows*).

2002-05-02d GBA

Er kíkvaðratpróf einhliða eða tvíhliða?

Þegar við reiknum kíkvaðrat próf erum við þá alltaf með einhliða próf og þar af leiðandi einungis líkindi á öðrum halanum eða er í einhverjum tilfellum reiknað tvíhliða próf?

Kíkvaðrat er ætíð tvíhliða próf í þeim skilningi að prófuð er tvíhliða tilgáta. Þegar niðurstöðu prófsins er flett upp í töflu er það meðhöndlað eins og það væri einhliða þar sem frávik frá núlltilgátunni leiða ætíð til hækkunar á kíkvaðrati.

Kíkvaðratpróf í krosstöflum er að þessu leytinu ólíkt t.d. t-prófi. Við samanburð tveggja meðaltala getur tilgátan ýmist haft stefnu eða verið stefnulaus. Ef tilgátan er stefnulaus, er horft fram hjá merkinu í niðurstöðum t-prófsins, þ.e. engu skiptir hvort niðurstaðan er jákvæð eða neikvæð og prófið sagt vera tvíhliða. Ef hins vegar tilgátan hefur stefnu, skiptir höfuðmáli hvort niðurstaðan er neikvæð eða jákvæð og því notað einhliða próf.

Í kíkvaðratprófi er niðurstaða prófsins alltaf jákvæð, þ.e. hærri en núll. Tilgátan er ætíð stefnulaus og tilgreinir það eitt að breyturnar tengist í þýði. Þrátt fyrir stefnulausa tilgátu er prófið sjálft hins vegar í reynd einhliða, því við reiknum líkindin á að fá svona háa eða hærri niðurstöðu í *plús*—hærri en núll—í stað þess að reikna líkindin á að fá svona mikið eða meira frávik í aðra hvora áttina, þ.e. í plús eða mínus.

2003-01-11i GBA

Væntitíðni í krosstöflum

Ég á í vandræðum með að reikna eitt verkefnið í stoðkverinu, hvernig reikna ég væntitíðni? Ég veit að það á að nota hlutfallið í þýðinu, sbr. glæru 1 um kíkvaðrat. En ég skil ekki almennilega formúluna fyrir væntitíðni, sbr. formúlublað. Ef þú gætir útskýrt þetta fyrir mér þá væri ég mjög þakklát.

Þú ættir að nota glæru 7 í staðinn. Við vitum ekki hvert hlutfallið er í þýðinu, það sem við getum vitað er hver hlutföllin (og þar með tíðnin) væru ef breyturnar væru óháðar.

Formúlan byggist á margföldunarreglunni en hún segir til um hverjar líkurnar eru á að tveir óháðir atburðir fari saman. Ef við skoðum glærunar Einfölduð tafla sjáum við að 26 af 255 nemendum eru í bókasafns- og upplýsingafræði eða um 10,2% nemenda. Sömuleiðis sjáum við að 84 af 255 nemendum eru karlar eða 32,9%. Margföldunarreglan segir að ef það að vera í bókasafns- og upplýsingafræði og það að vera karlmaður eru tveir óháðir atburðir, þá eru líkurnar $0{,}102 \cdot 0{,}329 = 0{,}034$ að saman fari það að vera karlmaður og að vera í bókasafns- og upplýsingafræði. Þessi líkindi (3,4%) samsvara því að 8,6 ($0{,}034 \cdot 255$) nemendur séu í senn karlar og í bókasafns- og upplýsingafræði og sú tala er því væntitíðnin fyrir þetta hólf töflunnar.

Lestu vel bls. 254–255 í Agresti og skoðaðu vel hvernig hann reiknar úr væntitíðni fyrir konur sem eru demókratar. Þú getur einnig skoðað umræðu um hvenær breytur eru óháðar á bls. 252–253. Sams konar umræða er á glærum 7 og 9. Þú getur spreytt þig við að reikna væntitíðni með því að reikna væntítíðnina á glæru 8 og bera saman við uppgefna væntitíðni.

2001-04-17a GBA

Túlkun staðlaðrar leifar

Glæra 11 í fyrirlestrinum um kíkvaðrat sýnir niðurstöður úr SPSS, m.a. leiðrétta leif. Ef ég skoða hana, þá sé ég að leiðrétt leif er jafnt og eða lægri en 2 í fjórum af fimm tilfellum. Er það að segja mér að í þessum fjórum tilfellum sé núll tilgátu ekki hafnað? Einnig væri gott að fá nokkur orð um hvað leiðrétt leif er.

Glæra 11 sýnir töflu með 10 hólfum. Við prófuðum þá tilgátu að aðalgrein til BA prófs (Námsgrein) tengdist kynferði nemanda. Núlltilgátan er því að kynferði og aðalgrein séu óháð. Núlltilgátuna prófuðum við með kíkvaðratprófi. Prófið var marktækt og því gátum hafnað núlltilgátunni og tekið upp aðaltilgátuna að kynferði tengist aðalgrein til BA prófs.

Leiðrétt leif hjálpar okkur við efnislega túlkun niðurstöðunnar, þ.e. að ákvarða hvers konar tengsl eru milli kynferðis og aðalgrein. Marktækt kíkvaðrat segir okkur aðeins að breyturnar Kynferði og Námsgrein tengist; við ályktum því að kynferði hafi áhrif á námsval. Það getur þýtt að ákveðnar greinar höfði meira til annars kynsins en að aðrar námsgreinar höfði jafnt til beggja kynja. Það er einnig hugsanlegt að flestar eða allar námsgreinar í félagsvísindadeild séu kyngreindar. Það að hafna núlltilgátunni segir í sjálfu sér ekkert um það hvernig þessi tengsl kynferðis og aðalgreinar til BA prófs eru nákvæmlega.

Leiðrétt leif getur veitt okkur mikilvægar upplýsingar um það hvers eðlis tengsl breytanna eru. Leiðrétt leif er mælikvarði á frávik frá núlltilgátunni (þ.e. mismun rauntíðni og væntitíðni) fyrir hvert hólf töflunnar fyrir sig. Ef leiðrétt leif er jákvæð (> 0) þýðir það að rauntíðni er hærri en væntitíðni; ef hún er neikvæð (< 0), er rauntíðni lægri en væntitíðni. Ef talnagildi leiðréttu leifarinnar er hátt (þ.e. mikið í mínus eða mikið í plús) þýðir það að mikill munur er á rauntíðni og væntitíðni (mikið frávik frá núlltilgátunni).

Til að meta hvað teljist hátt og hvað teljist lágt talnagildi fyrir leiðrétta leif má hafa í huga að leiðrétt leif er með sama mælikvarða og staðalvilla. Þannig eru 68% líkindi á því að fá leiðrétta leif á bilinu –1 til +1 og u.þ.b. 95% líkindi á að hún sé á bilinu –2 til +2. Því telst það óvenjulegt ef talnagildi leiðréttu leifarinnar er 2 eða hærra. En gættu þess að þetta er aðeins leiðsögn við túlkun niðurstaðna. Við erum að leita efnislegrar túlkunar í kjölfar þess að við höfnuðum núlltilgátunni, en ekki að prófa nýjar tilgátu fyrir hvert hólf töflunnar.

Skoðum nú glæru 11 í fyrirlestrinum um kíkvaðrat. Tölugildi leiðréttrar leifar er undir 2,0 í öllum hólfum nema fyrir bókasafnsfræði og stjórnmálafræði. Við drögum þá ályktun að lítill munur sé á raun- og væntitíðni fyrir sál-, félags- og mannfræði; kynjahlutfallið er samkvæmt því svipað og fyrir félagsvísindadeild í heild sinni.

Tölugildi leiðréttrar leifar er 2,0 fyrir bókasafnsfræði og 4,7 fyrir stjórnmálafræði. Þetta er óvanalega hátt og því ástæða til að skoða þær greinar nánar. Nánari samanburður á raun- og væntitíðni sýnir að í bókasafnsfræði eru heldur fleiri konur en búast mætti við og að í stjórnmálafræði eru tvöfalt fleiri karlar en búast mætti við. Efnislega túlkum við því niðurstöðurnar þannig að aðalgrein til BA prófs í félagsvísindadeild fari eftir kyni þannig að bókasafnsfræði laði frekar til sín konur en karla og stjórnmálafræði laði að sér karla umfram konur. Aðrar greinar virðast höfða til karla og kvenna í sömu hlutföllum og deildin í heild sinni.

1999-04-07b GBA

Styrkleiki tengsla þegar kíkvaðratpróf er ógilt

Ég hef verið að reikna og í nokkrum dæmanna í síðasta þriðjungi er kíkvaðratprófið ekki gilt. Samt er beðið um að reikna hversu sterk tengslin séu o.s.frv.

Ef prófið er ekki gilt þýðir þá nokkuð að halda áfram? Verður maður ekki fyrst að flokka saman eða taka út flokka og fá prófið gilt?

Ef prófið er ógilt hefur það engan tilgang að reikna marktekt. Því geturðu ekki fullyrt hvort sambandið (tengslin) sé til staðar eða ekki.

Eftir sem áður geturðu metið hversu sterkt sambandið er í úrtakinu, t.d. með τ, ϕ, o.s.frv. Það er ekkert í veginum; niðurstaðan vísar til styrkleikans í úrtakinu og tekur ekki afstöðu til þess hvort sambandið er fyrir hendi í þýði eða ekki.

Ef þér tekst að fá gilt kíkvaðratpróf með því að fella saman flokka, kemur auðvitað til álita að reikna sambandið annað hvort miðað við upprunalegu töfluna eða töfluna eftir að flokkar hafa verið felldir saman. Að sumu leyti er það réttara að nota upprunalegu töfluna en þó getur það verið álitamál. Hvor taflan er líklegri til að lýsa tengslunum í þýði? Því verður ekki svarað hreint út heldur verður að meta það í hvert skipti.

Ef síðan kíkvaðrat er gilt en ómarktækt, kemur til álita að reikna ekki styrkleika tengslanna. Í slíkum kringumstæðum ályktar þú að tengslin séu ekki fyrir hendi í þýði og því hálfgert vindhögg að meta styrkleika þeirra. Ekkert bannar þér þó að gera það og í prófi myndi ég ekki hika við að gera það sem mér er sagt.

2000-04-29c GBA

Ef svona kemur í prófi, hvort á þá a) að halda áfram og reikna tengslin eða b) segja að prófið sé ekki gilt og koma með útskýringuna á af hverju og hvað beri að gera til að fá prófið gilt?

Ég held ég hafi svarað þessu óbeint. Ég myndi taka þau skref sem þyrfti til að gera prófið gilt. Síðan myndi ég reikna styrkleika tengslanna miðað við annað hvort upprunalegu eða breyttu töfluna. Síðan myndi ég setja útskýringu á því hvers vegna ég fer þessa leið til að verja mig fyrir vondum kennurum; slíkir kennarar eiga ekki að finnast í Aðferðafræði II en allur er varinn góður.

2000-04-29c GBA

Hvað er skilyrt dreifing?

Hver er munurinn á skilyrtri og óskilyrtri dreifingu?

Gera má greinarmun á skilyrtri (*conditional distribution*) og jaðardreifingu (*marginal distribution*). Jaðardreifing er dreifing breytu beint af skepnunni ef svo má segja. Skilyrt dreifing er hins vegar dreifing breytunnar (oftast fylgibreytu) þegar einhver önnur breyta (oftast frumbreyta) tekur einhver tiltekin gildi.

Ég get hugsað mér að líkamshæð dreifist ákveðið mikið og metið dreifinguna t.d. með staðalfráviki. Þetta væri þá jaðardreifing líkamshæðar. Ég get skilyrt þessa dreifingu á ýmsan máta. Ég get skoðað dreifingu líkamshæðar eingöngu fyrir konur. Þá hef ég skilyrt dreifinguna þannig að kynferðið sé kona. Á sama hátt get ég skoðað dreifingu hjá körlum en það væri þá önnur skilyrt dreifing, þ.e. líkamshæðardreifing karla.

Þú sérð dæmi um skilyrta dreifingu af þessu tagi í fyrirlestrinum Sveigfylgni, þ.e. skilyrta dreifingu þar sem frumbreytan er eigindleg en fylgibreytan megindleg.

Hugtökin geta skýrst ef við skoðum krosstöflur. Taflan til hliðar sýnir hversu margir íslenskir og erlendir dagskrárliðir voru í fjórum íslenskum sjónvarpsstöðvum samkvæmt lauslegri talningu einn föstudag að vori. Heildarskipting allra dagskrárliða sést lengst til hægri (í jaðrinum) á töflunni; þetta er jaðardreifingin enda í jaðri töflunnar. Þar sést að vel innan við helmingur dagskrárliða eru íslenskir og um 60% þeirra eru erlendir.

Tegund efnis	Sjón varpið	Stöð 2	Sýn	Skjár einn	Samtals
Íslenskt	9	7	3	6	25
Erlent	5	21	6	7	39
Samtals	14	28	9	13	64

Ef við skoðum skiptingu efnis aðeins fyrir Sjónvarpið sést að þar eru íslenskir dagskrárliðir í meirihluta eða um 64%. Hér erum við að skoða skilyrta dreifingu, þ.e. dreifingu í tegund efnis þegar við skoðum aðeins efni hjá Sjónvarpinu. Við getum einnig skoðað skiptinguna hjá Stöð 2; þar er enn önnur skilyrt dreifing. Þar er erlendir dagskrárliðir í miklum meirihluta eða um 75%. Þannig getum við haldið áfram og skoðað dreifinguna fyrir hverja sjónvarpsstöð fyrir sig. Við erum því með fjórar skilyrtar dreifingar í töflunni, eina fyrir hverja sjónvarpsstöð.

Það má líka tala um skilyrta dreifingu í tengslum við aðfallsgreiningu. Þar fáum við línu sem gefur bestu spágildi fylgibreytunnar fyrir hvert og eitt gildi frumbreytunnar. Raungildin dreifast í kringum línuna. Setjum nú sem svo að línan lýsi þyngd eftir líkamshæð. Þá get ég spurt hver líkamsþyngd fólks sé sem er 178 cm á hæð. Línan og aðfallsjafnan gæfi mér besta spágildið. Hins vegar munu fáir hafa nákvæmlega sömu líkamsþyngd og spágildið gefur til kynna; sumir eru léttari en aðrir þyngri. Þannig dreifast einstaklingarnir í kringum spágildið. Við getum því spurt: Hver er þyngdardreifing þeirra sem eru 178 cm á hæð? Sú dreifing samsvarar dreifingunni í kringum línuna og væri skilyrt dreifing fyrir 178 cm hátt fólk. Jaðardreifingin væri hins vegar heildardreifing líkamsþyngdar, þ.e. þyngdardreifing allra óháð líkamshæð.

Ég er því búinn að nefna þrjú dæmi um skilyrtar og jaðardreifingar. Fyrsta lagi þar sem frumbreytan er eigindleg en fylgibreytan megindleg, í öðru lagi þar sem báðar breytur eru eigindlegar og að síðustu þar sem báðar breytur eru megindlegar.

Það er fjallað um skilyrta og jaðardreifingar í fyrirlestrunum Mælitölur á tengsl breyta og Kíkvaðrat. Í báðum fyrirlestrum finnurðu glærur sem fjalla beinlínis um þessi hugtök. Skoðaðu þær vel í tengslum við þessa spurningu.

2000-04-28 GBA

Kíkvaðrat, væntitíðni og α

Á ég ekki að sjá það á töflunni þegar ég er með kíkvaðrat hvort væntitíðnin sé minni en 5 í 20% reitanna?

Þú þarft að reikna væntitíðnina fyrir hvert hólf fyrir sig, þetta er fyrsta skrefið í því að reikna kíkvaðratpróf fyrir krosstöflu. Síðan þarftu að telja hversu oft væntitíðnin er undir 5 og sjá þannig hvort það eru meira en 20% hólfanna í töflunni.

Ég ekki alveg viss hvernig ég á að finna $p < 0{,}01$ eða $p > 0{,}01$ fyrir kíkvaðrat.

Þú færð niðurstöðu úr kíkvaðratprófinu, prófniðurstöðuna. Ef hún er há, er frávikið frá núlltilgátunni mikið. Þú vilt að niðurstaðan sé það há að það séu minna en 5% líkur á því að fá þetta mikið eða meira frávik þegar núlltilgátan er rétt. Þú sérð hversu há prófniðurstaðan þarf að vera með því að fletta upp í kíkvaðrattöflu. Það er gefið sýnishorn af slíkri töflu í fyrirlestrinum Kíkvaðrat.

Athugaðu að það er ekkert sem segir að þú eigir að miða við $\alpha = 0{,}01$; algengast er að miða við $\alpha = 0{,}05$.

Er ekki rétt að athuga skilyrtu dreifinguna til að athuga með tengsl milli breyta í dæmi 92 í Stoðkveri?

Jú, en athugun á skilyrtu dreifingunum er aðeins ein af nokkrum leiðum sem þú getur notað. Fleiri möguleikar eru nefndir í fyrirlestrinum Kíkvaðrat.

2003-04-30a GBA
2003-12-29b GBA

Fyrir hvað stendur ρ?

Hvað er $p = 0{,}0$ á glæru 2 í **Afköst rannsókna**? Er þetta bara venjulegt p-gildi eða eitthvað annað?

Þetta er ekki *p* heldur gríski stafurinn ró! Grískir stafir eru notaðir fyrir þýðistölur og rómverskir fyrir úrtakstölur. Pearson fylgni í úrtaki er því táknað með *r* en samsvarandi fylgni í þýði er táknað með ρ.

Á þessari tilteknu glæru táknar þetta því að gert er ráð fyrir að engin fylgni sé í þýði. Síðan eru endurtekið dregið 50 manna úrtak úr þessu þýði. Einstaklingar veljast úr þýðinu í úrtakið samkvæmt tilviljun. Fylgnin í úrtakinu verður því stundum hærri en 0,0 og stundum lægri en 0,0 (þ.e. neikvæð) vegna þess að einstaklingarnir raðast í úrtakið þannig að úrtaksfylgnin verður ekki nákvæmlega 0,0. Ef þú hugleiðir það, sérðu að það er mjög erfitt að velja 50 einstaklinga úr þýði þannig að fylgnin í þessu 50 manna úrtaki endurspegli nákvæmlega fylgnina í þýðinu.

Fylgnin í þýðinu er sem sé 0,0, þ.e. $\rho = 0{,}0$. Fylgnin í úrtakinu verður ýmist hærri eða lægri; myndin sýnir þessa dreifingu á Pearson *r* frá einu úrtaki til annars.

2001-04-27a GBA

Hvernig eyk ég áhrif í þýði?

Ég er búin að velta afköstum mikið fyrir mér en skil þau þó ekki enn að fullu. Ég skil vel að það er gott að velja stærra α (t.d. 0,05 í stað 0,01) til þess að ná meiri afköstum. En ég fatta hins vegar ekki hvernig maður eykur áhrifin í þýðinu. Eykur maður áhrifin í þýðinu til þess að hafa meðaltal undir aðaltilgátu sem ólíkast þýðismeðaltali undir núlltilgátu?

Einnig fatta ég ekki hvað felst í því að auka áhrif með því að stjórna villuþáttum.

Það er rétt að aukin áhrif í þýði felst í því að þýðismeðaltalið undir aðaltilgátunni verður ólíkara þýðismeðaltalinu undir núlltilgátunni eykst; frávikið frá núlltilgátunni eykst því. Afleiðing þess er að úrtaksmeðaltalið verður að jafnaði ólíkara viðmiðsgildinu (þýðismeðaltalinu samkvæmt núlltilgátunni) og því verður tölfræðiprófið oftar marktækt.

Áhrif í þýði get ég fyrst og fremst aukið með réttri hönnun rannsóknarinnar. Ef ég er með tilraun, get ég aukið áhrifin með því að auka styrk inngripsins (*intervention*). Í fylgnirannsókn, leita ég að kringumstæðum þar sem dreifisvið frumbreytunnar er sem mest. Til að skoða hvort námsframmistaða nemenda með góðar námsaðferðir sé betri en frammistaða nemenda með slakar námsaðferðir, myndi ég vilja skoða þýði þar sem mikill munur er á námsaðferðum. Þannig væri betra að skoða nemendur í 10. bekk grunnskóla eða 1. misseri í framhaldsskóla heldur en nemendur í Háskóla, því að öllum líkindum fækkar þeim sem er með allra verstu námsaðferðirnar eftir því sem skólastigið er hærra.

Afköstin ráðast einnig af staðalvillunni, því minni sem staðalvillan er því meiri afköst svo fremi sem áhrifin minnki ekki að sama skapi. Við höfum lagt áherslu á framlag úrtaksvillunnar (*sampling error*) til úrtakadreifingarinnar. En hluti úrtaksvillunnar er tilkomin vegna ónákvæmni í mælingu. Ef ég bæti mælinguna mun staðalfrávikið og því staðalvillan minnka sem því nemur án þess að áhrifin í þýði breytist. Því getur vandvirkni í skipulagningu og framkvæmd rannsóknarinnar orðið til þess að auka afköst hennar.

2001-03-15a GBA

Hvað er fullnægjandi nákvæmni?

Á glæru 7 í fyrirlestrinum **Úrtak sem hlutfall af þýði** segir að fullnægjandi nákvæmni fáist ef útakið er af stærð 28. Er þá miðað við að úrtak án skila sé það stórt að öryggisbil sé ekki breiðara en 10 prósentustig? Er það almennt viðmið fyrir fullnægjandi nákvæmni? Hvaða viðmið nota ég þá á glæru 8?

Það er ekki hægt að gefa neina reglu í þessu efni, nákvæmin fer eftir því sem þú leitar hverju sinni. Aukin nákvæmni felur yfirleitt í sér stærra úrtak og meiri kostnað í peningum, mannafla eða tíma. Best er að hafa fullkomna nákvæmni en þessi kostnaðarauki kemur í veg fyrir slíkt.

Á glærunni **Staðalvilla eftir hlutfalli í þýði** er ætlunin að meta hlutfall sem búist er við að sé einhvers staðar nálægt 0,15 í þýði. Augljóslega þarf að staðsetja þetta lágt hlutfall með tiltölulega mikilli nákvæmni. Ef ég ætlaði mér t.d. að láta mér nægja að spyrja 16 kennara, gæti ég búist við því að staðalvillan yrði 0,063 og frávikið gæti því verið næstum 15 prósentustig í hvora átt miðað við 95% öryggi. Augljóslega er þetta of lítil nákvæmni. [Sem þumalputtareglu getur þú tvöfaldað staðalvilluna til að fá 95% vikmörk. Þar sem úrtakið er þetta lítið vanmetur það öryggisbilið; með því að fletta í t-töflu gætirðu fengið nákvæmari niðurstöðu.]

Með því að stækka úrtakið minnkar staðalvillan. Fljótt á litið virðist sem úrtak með 28 kennurum gefi nægjanlega nákvæmni. Staðalvillan væri 0,024 og 95% öryggisbil því með u.þ.b. 5 prósentustiga vikmörk í hvora átt. Við gætum því staðsett þýðishlutfall innan bils sem væri 10 prósentustig að vídd. Hvort það er nægilega nákvæmt fer eftir því hvað þú ætlar þér með upplýsingarnar en fljótt á litið virðist sem þú myndir ekki sætta þig við minni nákvæmni en þetta. Ef ég hins vegar vel 28 kennara í úrtakið, þá gæti ég allt eins tekið alla og fengið þannig upplýsingar um þýðistölurnar sjálfar.

Á glærunni **Laun sálfræðinema eftir útskrift** er sama upp á teningnum; ég get ekki sagt þér hversu mikla nákvæmni þú ættir að leita eftir. Fljótt á litið virðist þó að þú myndir vilja a.m.k. 71 nemanda í úrtakið þar sem það gefur öryggisbil sem er 20.000 króna breitt ($2 \cdot 1{,}96 \cdot 4.740 = 18.581$). Öryggisbilið yrði helmingi minna ef þú notaðir 160 nemendur og þá gætirðu staðsett meðallaunin af tiltölulega mikilli nákvæmni. Ef þú vilt enn meiri nákvæmni, þarftu enn að stækka úrtakið. Á endanum þarftu sjálf að ákveða miðað við þau not sem þú vilt hafa af upplýsingunum og þann kostnað sem þú getur borið af því að afla þeirra.

2001-04-03a GBA