Aðferðafræði II 10.05.03

Spurt og svarað

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

II. Öryggisbil og ályktanir í einum hópi

Fræðilíkur og raunlíkur

Fræðilíkurnar á að fá landvættina þegar maður kastar upp peningi er ½ en eru raunlíkurnar það ekki líka þegar til lengri tíma er litið? Ef svo er hver er þá munurinn á þessu tvennu?

Jú, þetta er rétt. Munurinn er sá að fræðilíkurnar eru alltaf þær sömu en raunlíkurnar eru breytilegar og óvissar. Að jafnaði þegar til lengdar er litið nálgast þær fræðilíkurnar því meir sem tilraununum fjölgar. Þetta miðast þó auðvitað við það að fræðilíkurnar séu rétt reiknaðar, þ.e. að líkindalíkanið sé rétt.

(2002-03-11b GBA)
2003-01-11c GBA

Líkindi á einum atburði í nokkrum tilraunum

Mig langar mikið til að fá aðeins hjálp við að skilja betur líkindareikninginn. Hvernig reikna ég líkindin fyrir því að atburður gerist t.d. 2 í þremur tilraunum eða eitthvað svoleiðis.

Setjum sem svo að fimmtungur þeirra bifreiða sem eru á ferðinni á rúmhelgum degi í Reykjavík sé jeppi. Þá mætti búa til nokkur dæmi á þeim grundvelli. Ég gæti til dæmis spurt (a) hversu líklegt sé að mæta þremur jeppum í röð, (b) hversu líklegt sé að mæta þremur bílum og að minnsta kosti tveir séu jeppar og (c) hversu líklegt sé að mæta þremur bílum og enginn sé jeppi. Þetta eru auðvitað aðeins þrír möguleikar og fjölmörg önnur dæmi væri hægt að búa til.

Fyrst skulum við telja upp möguleikana. Ég nota „J“ fyrir atburðinn að mæta jeppa og „¬J“ fyrir það að mæta ekki jeppa. Líkurnar á því að mæta jeppa er P(J)=

0,20 (þ.e. fimmtungur bíla eru jeppar) og líkurnar á að mæta ekki jeppa eru þá P(¬J)= 1 - 0,20= 0,80. Þá eru eftirfarandi möguleikar:

Atburður	**Líkindi**
1. J, J, J	P(J, J, J)= P(J) · P(J) · P(J)= 0,2 · 0,2 · 0,2= 0,008
2. J, J, ¬J	P(J, J, ¬J)= P(J) · P(J) · P(¬J)= 0,2 · 0,2 · 0,8= 0,032
3. J, ¬J, J	P(J, ¬J, J)= P(J) · P(¬J) · P(J)= 0,2 · 0,8 · 0,2= 0,032
4. J, ¬J, ¬J	P(J, ¬J, ¬J)= P(J) · P(¬J) · P(¬J)= 0,2 · 0,8 · 0,8= 0,128
5. ¬J, J, J	P(¬J, J, J)= P(¬J) · P(J) · P(J)= 0,8 · 0,2 · 0,2= 0,032
6. ¬J, J, ¬J	P(¬J, J, ¬J)= P(¬J) · P(J) · P(¬J)= 0,8 · 0,2 · 0,8= 0,128
7. ¬J, ¬J, J	P(¬J, ¬J, J)= P(¬J) · P(¬J) · P(J)= 0,8 · 0,8 · 0,2= 0,128
8. ¬J, ¬J, ¬J	P(¬J, ¬J, ¬J)= P(¬J) · P(¬J) · P(¬J)= 0,8 · 0,8 · 0,8= 0,512

Auðvitað þarftu ekki að reikna alla þessa atburði til að svara spurningunum þremur, heldur sýni ég þessa útreikninga til þess að sýna hvernig hægt er að leysa svona dæmi með því að telja upp alla möguleika og nota margföldunar- og samlagningarreglurnar. Í reynd nægir okkur að reikna líkindin á atburðum 1, 4, 6, 7 og 8, sbr. útreikninganna hér fyrir neðan.

Þá eru það dæmin þrjú.

Byrjum á lið (a): Hversu líklegt er að mæta þremur jeppum í röð. Þetta er nákvæmlega atburður nr. 1 hér fyrir ofan:

$$P(J, J, J)= P(J)\cdot P(J)\cdot P(J)= 0{,}2\cdot 0{,}2\cdot 0{,}2= 0{,}008$$

Liður (b): Hverjar eru líkurnar á *að minnsta kosti* tveir bílar af þremur séu jeppar? Þetta samsvarar því að annað hvort tveir bílar eða þrír bílar af þremur séu jeppar. Það eru þrír mismunandi möguleikar á að mæta tveimur jeppum og einum bíl sem er ekki jeppi. Þetta samsvarar atburðum nr. 2, 3 og 5 hér að ofan. Það er aðeins einn möguleiki á að mæta þremur jeppum, þ.e.a.s. atburður 1.

P(að minnsta kosti tveir jeppar) = P(einn jeppi og tveir ekki jeppar) + P(þrír jeppar) = P(J, J, ¬J) + P(J, ¬J, J) + P(¬J, J, J) + P(J, J, J = 0,032 + 0,032 + 0,032 + 0,008 = 0,104

Liður (c): Hverjar eru líkurnar á að mæta þremur bílum og enginn sé jeppi. Þetta samsvarar atburði nr. 8 hér fyrir ofan:

$$P(\neg J, \neg J, \neg J)= P(\neg J)\cdot P(\neg J)\cdot P(\neg J)= 0{,}8\cdot 0{,}8\cdot 0{,}8= 0{,}512$$

Allir byggja þessir útreikningar á því að þetta séu óháðir atburðir, þ.e. að ef ég mæti einum jeppa þá hvorki auki það né minnki líkindi þess að næsti bíll verði jeppi. Þetta ætti að standast nokkurn veginn á venjulegum degi, en til dæmis alls ekki á jeppadegi Toyotaumboðsins.

1999-03-24 GBA
1999-04-08 GBA

Margföldunarreglan og samsett líkindi ‡

Ég átta mig ekki á glærunni um margföldunarregluna í líkindafræðinni (Líkindafræði: Margföldunarregla). Ef ég skil hana rétt þá eru líkindin á því að fá 2 *og* 4 í tveimur teningsköstum 1/36.

Á næstu glæru (Að fá 2 *og síðan* 4) er sagt að líkurnar á að fá 2 í fyrsta kasti og 4 í því seinna sé 1/36. Þetta eru sömu líkindin á báðum glærum en ólíkir atburðir.

Eru ekki tveir möguleikar á að fá 2 og 4 í tveimur köstum? Ég get fengið fyrst 2 og síðan 4 eða fyrst 4 og síðan 2, er það ekki? Þá hljóta líkindin að eiga að vera tvöfalt meiri, þ.e. 2/36.

Geturðu staðfest að ég hugsi þetta rétt?

Það er rétt að þessar glærur bjóða upp á misskilning og fyrri glæran er ónákvæm. Ég tel mig hafa útskýrt glæruna þannig að um sama atburð sé að ræða á báðum glærum, þannig eru a.m.k. fyrirlestrarglósurnar mínar. En það er mikilvægt að glærurnar séu skýrar og geti staðið sjálfstætt. Ég hef þess vegna breytt orðalaginu á fyrri glærunni svo það verði eins og á þeirri seinni. Glærurnar eru því réttar eins og þær eru fyrir utan ónákvæmt orðalag í þeirri fyrri, sem ég hef lagað.

Þú spyrð einnig um líkur þess að fá bæði 2 og 4 í tveimur köstum. Þú reiknar þau líkindi alveg rétt en ég læt samt nákvæmari útlistun fljóta með.

Eins og þú segir eru tveir möguleikar á að fá bæði 2 og 4 í tveimur teningsköstum. Ég get fengið fyrst 2 og síðan 4 og líkur þess eru 1/36 ($P(2;\ 4) = P(2)\cdot P(4) = 1/6\cdot 1/6 = 1/36$). Ég get líka fengið fyrst 4 og síðan 2 og líkur þess eru einnig 1/36 ($P(4;\ 2) = P(4)\cdot P(2) = 1/6\cdot 1/6 = 1/36$). Þetta eru því tveir atburðir og líkindi hvors atburðar fyrir sig er 1/36.

Líkindi þess að fá bæði 2 og 4 í tveimur köstum er því líkur þess að fá 2 og síðan 4 eða að fá 4 og síðan 2. Heildarlíkurnar fáum við með samlagningarreglunni. Við getum táknað reiknað þetta svona: P(2 og 4 í tveimur köstum)= P([2 og síðan 4] eða [4 og síðan 2])= P(2 og síðan 4) + P(4 og síðan 2) = $P(2)\cdot P(4) + P(4)\cdot P(2) = 1/6\cdot 1/6 + 1/6\cdot 1/6 = 1/36 + 1/36 = 2/36 = 1/18 \approx 0{,}056$.

Hvað greinir á milli annars vegar þess að fá 2 og síðan 4 og hins vegar þess að fá bæði 2 og 4 í tveimur köstum? Hið fyrra er einn atburður samsettur úr tveimur óháðum samrýmanlegum frumatburðum og því gildir margföldunarreglan. Hið síðara eru tveir ósamrýmanlegir atburðir sem hvor um sig er samsettur úr frumatburðum; því þarf að nota bæði margföldunar- og samlagningarregluna. Það að fá fyrst 2 og síðan 4 útilokar það að fá 4 og svo 2; þetta eru því ósamrýmanlegir atburðir. Það að fá 2 samrýmist því að fá 4 svo og öfugt; frumatburðirnir eru því samrýmanlegir.

2000-03-22 GBA

Endanlegt og óendanlegt þýði

Hver er munurinn á endanlegu og óendanlegu þýði? Geturðu nefnt dæmi um hvort tveggja svo ég fái einhverja hugmynd í hausinn?

Nemendur í Aðferðafræði II eru endanlegt þýði, við getum talið þá. Allir jarðarbúar eru líka endanleg stærð og því endanlegt þýði.

Ef ég hins vegar vill vita hversu hátt bolti skoppar (að meðaltali) sem ég læt detta niður af þriðju hæð, þá gæti ég hent boltanum óendanlega oft (auðvitað mun hann þó eyðileggjast að lokum). Þýði boltakasta er því óendanlegt. Ég get komist að þessum eiginleika boltans (skopphæfni hans) með því að taka úrtak úr þessu þýði, þ.e. hent honum niður nokkrum sinnum. Á grundvelli þeirra upplýsinga get ég ályktað um þýðið, þ.e. meðaltal allra hugsanlegra boltakasta, sem samsvarar skopphæfninni sjálfri.

Vor 2000 GBA

Hver er munurinn á σ og $\hat{\sigma}$?

Er það ekki rétt að staðalfrávik úrtaks sé ýmist táknað með s eða $\hat{\sigma}$, þ.e. er ekki örugglega um það sama að ræða í báðum tilfellum?

Nei, það er ekki um nákvæmlega það sama að ræða en talnalega niðurstaðan er svipuð. Mælitalan s er úrtaksstaðalfrávikið en $\hat{\sigma}$ er spágildi (spátala) fyrir þýðis-staðalfrávikið.

Við notum oftast samsvarandi úrtakstölu sem spátölu fyrir viðkomandi þýðistölu, en það er ekkert sem segir að svo sé í öllum tilvikum. Þetta eru því tvö ólík hugtök sem vert er að halda aðskildum. Úrtaksstaðalfrávikið s er t.d. skekkt mælitala og því ekki hentug sem spátala fyrir þýðisstaðalfrávikið. Ef hins vegar deilt er með N-1 í nefnaranum í stað N eins og í formúlunni fyrir s, fæst óskekkt spátala fyrir þýðisstaðalfrávikið. Þegar unnið er með staðalfrávik er það því sú mælitala sem er notuð sem $\hat{\sigma}$.

Á formúlublaðinu er formúlan fyrir s gefin upp með N í nefnaranum en spágildið ($\hat{\sigma}$) með N–1 í nefnaranum. Þetta er í samræmi við ofangreint: Ef nota á staðalfrá-vikið eingöngu til að lýsa dreifingu úrtaksins er rétt að nota N í nefnaranum. Ef nota á staðalfrávikið sem spágildi, þ.e. sem $\hat{\sigma}$, þarf að deila með N–1 til að fá óskekkta spátölu.

Auðvitað *gætum* við notað niðurstöðuna fyrir s sem spátölu en þá fengjum við $\hat{\sigma}$ sem væri skekkt niður á við. Það er afar óæskilegt og því er spátalan fyrir þýðis-staðalfrávik jafnan reiknað eins og formúlublaðið gefur til kynna.

2003-01-11f GBA

Skekkja og villa

Er það rétt skilið hjá mér að skekkja sé sveifla kringum „ranga" gildið, þ.e. að annað hvort er það réttur staður eða ekki, og að villa sé sveifla kringum rétta gildið vegna tilviljunarkennds fráviks í úrtökum?

Ja, ekki alveg en í aðalatriðum rétt. Ef mælitalan er óskekkt, er villan breytileiki (sveifla) í kringum þýðistöluna (rétta gildið) en annars felst villan í breytileika í kringum ranga (skekkta) gildið. Skekkja er það kallað ef úrtakstalan (niðurstaðan í úrtakinu) er að jafnaði önnur en þýðistalan.

Leyfðu mér að kynna nýtt hugtak, væntigildi (expected value) úrtakstölunnar. Væntigildið er þá meðaltal úrtakstölunnar reiknað yfir óendanlega mörg úrtök. Skekkjan er þá fjarlægð væntigildisins frá þýðistölunni, þ.e. hve miklu munar að jafnaði á úrtakstölunni og þýðistölunni. Ef mælitala er óskekkt, er væntigildið jafnt þýðistölunni. Ef hún er skekkt, er einhver munur—mikill eða lítill—á þýðistölunni og væntigildinu.

Athugaðu að ef úrtakstalan er skekkt gefur hún að jafnaði ranga niðurstöðu. Eftir sem áður verður niðurstaðan ólík eftir úrtökum og því verður breytileiki frá einu úrtaki til annars. Þessi breytileiki er villa; villan er í kringum *ranga* gildið í þessu tilviki. Það væri þá bæði skekkja (kerfisbundið frávik frá rétta gildinu) og villa (tilviljunarkennd frávik frá væntigildinu) í niðurstöðunum.

Villa og skekkja eru því óháð (en skyld) hugtök; allar niðurstöður í úrtaki eru með villu en sumar einnig með skekkju.

Ef við notum skekkt staðalfrávik vanmetum við þá alltaf breytileika þýðis eða er það misjafnt hvort við van- eða ofmetum?

Óleiðrétt staðalfrávik (þ.e. deilt með N en ekki N–1) er skekkt niður á við. Þessi skekkja er mest í litlum úrtökum. Að jafnaði er því um vanmat að ræða, en vegna villunnar verður ofmat í sumum tilvikum. Ef skekkjan væri engin væri ofmat jafn algengt og vanmat. Í stórum úrtökum er skekkjan lítil og vanmat því aðeins lítillega algengara en ofmat. Í litlum úrtökum verður skekkjan meiri og vanmat því mun algengara en ofmat.

(2002-03-11b GBA)
2003-01-11d GBA
2003-12-29f GBA

Meðaltal meðaltala

Hvernig finn ég meðaltal meðaltalanna?

Í Aðferðafræði II ræðum við um meðaltal meðaltala í sambandi við óskekktar (*unbiased*) mælitölur og úrtakadreifingu (*sampling distribution*). Meðaltal meðaltalna reiknum (finnum) við aldrei nema ef ske kynni að við værum að vinna með niðurstöður hermilíkans.

Ef við drögum fjöldamörg tilviljunarúrtök úr þýði og reiknum óskekkta mælitölu í hverju og einu þeirra, þá gefur sú tala óskekkt spágildi fyrir samsvarandi þýðistölu. Lykileiginleiki slíkrar óskekktrar spátölu er að ef meðaltal hennar er reiknað yfir mörg tilviljunarúrtök nálgast það meðaltal samsvarandi þýðistölu meir og meir eftir því sem fjöldi úrtaka eykst. Ef fjöldi úrtaka er óendanlega mikill, verður meðaltal spátalnanna (yfir þessi óendanlega mörgu úrtök) nákvæmlega jafnt þýðistölunni.

Í sambandi við úrtaksmeðaltöl þýðir þetta að ef fjöldamörg tilviljunarúrtök eru dregin úr þýði og úrtaksmeðaltal sérhvers þeirra er reiknað, þá gefur meðaltal allra þessara úrtaksmeðaltalna allnákvæma mynd af þýðismeðaltalinu og því nákvæmari sem úrtökin verða fleiri. Ef úrtökin eru óendanlega mörg, verður þetta meðaltal úrtaksmeðaltala nákvæmlega jafnt þýðismeðaltalinu.

2001-04-10b GBA

Munurinn á staðalvillu og staðalfráviki

Hver er munurinn á staðalvillu og staðlafráviki?

Staðalfrávik er mælitala á breytileika í úrtaki eða þýði en staðalvilla er mælitala á breytileika mælitölu (t.d. meðaltals) frá einu úrtaki til annars.

Staðalfrávik er mælitala á breytileika (dreifingu) einstaklinga eða staka. Ef við fáum t.d. upplýsingar um sumartekjur nemenda sem stunda nám við félagsvísindadeild, þá getum við reiknað staðalfrávik teknanna. Ef staðalfrávikið er hátt, eru tekjur nemenda mjög ólíkar (breytilegar); ef það er lágt eru tekjur mjög svipaðar, þ.e. flestir með tekjur á tiltölulega þröngu tekjubili. Ef tekjurnar eru normaldreifðar getum við túlkað staðalfrávikið þannig að 68% hafi tekjur á bili sem er frá einu staðalfráviki fyrir neðan og upp í eitt staðalfrávik fyrir ofan meðaltal allra nemendanna.

Staðalfrávik getur hvort sem er vísað til þýðis eða úrtaks. Þannig get ég reiknað staðalfrávik sumartekna fyrir alla nemendur í félagsvísindadeild (þýðið) eða aðeins hluta þeirra (t.d. 100 manna tilviljunarúrtak nemenda í félagsvísindadeild).

Staðalvilla vísar til breytileika frá einu úrtaki til annars. Ef ég dreg endurtekið tilviljunarúrtak úr nemendum deildarinnar og reikna (t.d.) meðaltal fyrir hvert úrtak, þá verða meðaltölin jafn mörg og ólík og úrtökin eru mörg. Ef úrtökin er stórt, t.d. 500 nemenda úrtök, verða öll meðaltölin þó áþekk, þ.e. meðaltölin munu hafa lítinn breytileika. Ef úrtökin eru lítil, t.d. 5 nemenda úrtök, verða meðaltölin mjög ólík (breytileg) frá einu úrtaki til annars. Staðalvillan er mælikvarði á þennan breytileika milli úrtaka, svonefnda úrtakadreifingu (sampling distribution).

Staðalvillu get ég ýmist metið með því að draga fjöldamörg úrtök úr sama þýði og meta breytileika milli úrtaka eða nota viðeigandi formúlur. Hið fyrra er óframkvæmanlegt í flestum tilfellum en ef eiginleikar þýðis eru þekktir má líkja eftir þessu með svokölluðum Monte-Carlo tilraunum, þ.e. tölvulíkani af þýði og forritum sem draga fjölmörg úrtök á grunni slíkra líkana. Formúlur eru hins vegar til fyrir staðalvillur fjölmargra mælitalna. Þannig má t.d. reikna staðalvillu meðaltals á grundvelli úrtaksstærðar og staðalfráviks í úrtaki eða þýði. Fjölmargar aðrar mælitölur hafa staðalvillur, t.d. hallastuðlar (aðhvarfsstuðlar) og spágildi (*estimate*) í aðfallsgreiningu, þótt ekki sé fjallað um þær í Aðferðafræði II.

1999-04-07 GBA

Úrtakadreifing og dreifing í úrtaki

Ég skil ekki hvernig ég á að reikna dæmi 4.31 í Agresti og finn það hvergi í bókinni. Ég býst við að þetta snúist um aðferð til að finna líkurnar á að eitthvað sé ákveðið mörg staðalfrávik frá meðaltali, en hvaða formúla er fyrir þetta?

Í sjálfu sér þarf enga formúlu en ef þú vilt geturðu notað formúluna fyrir *z*-gildi úr Aðferðafræði I og formúluna fyrir *z*-próf úr Aðferðafræði II. Síðan þarf að fletta upp í normaltöflu, t.d. töflu A í Agresti.

Þýðismeðaltalið er \$250 og þýðisstaðalfrávikið \$75. Þú vilt finna hversu líklegt gildi upp á \$300 eða hærra miðað við að þýðið sé normaldreift. Þetta gildi er \$50 yfir þýðismeðaltalinu en það samsvarar 50/75= 2/3 ≈ 0,67 staðalfrávikum fyrir ofan meðaltalið. Ef ég fletti *z*-töluna 0,67 upp í töflu A í Agresti fæ ég hlutfallið 0,2514, þ.e. 25% verkamannanna er með 300 dollara eða meira í vikulaun.

Til að finna líkindi þess að 9 manna hópur hafi að meðaltali 300 dollara í vikulaun eða meira þarf ég að komast að því hver úrtakadreifing (*sampling distribution*) úrtaksmeðaltalanna er. Ég veit staðalfrávikið í þýðinu svo ég get reiknað staðalvillu meðaltalanna og miðað við normaldreifingu (en ekki t.d. *t*-dreifingu).

Staðalvillan er staðalfrávik þýðis deild með kvaðratrótinni af úrtaksstærðinni. Þetta samsvarar því 75/3 eða 25. Meðaltalið er 50 dollurum frá þýðismeðaltalinu en það samsvarar því að úrtaksmeðaltalið sé 2 (50/25) staðalvillum fyrir ofan. Ef ég fletti 2,0 upp í töflu A í Agresti sé ég að 0,0228 af svæðinu undir normalferlinum er þar fyrir ofan. Því álykta ég að rétt rúm 2% úrtaksmeðaltala í 9 manna úrtökum séu \$300 eða hærri, þ.e. líkurnar á svona háu eða hærra meðaltali er rétt rúm 2%.

Þetta dæmi er gott fyrir þá sök að það hjálpar að greina á milli dreifingu í úrtaki (*sample distribution*) og úrtakadreifingu (*sampling distribution*).

1999-03-05 GBA

Hver er munurinn á „sample distribution“ og „sampling distribution“ á einfaldri íslensku.

Enska orðasambandið „*sample distribution*“ er þýtt sem dreifing í úrtakinu, þ.e. dreifing staka í úrtaki. Við höfum mæligildin Y_1, Y_2, Y_3... til Y_n sem hafa dreifingu og miðsækni. Dreifingu í úrtakinu má skoða myndrænt til dæmis með stöplariti eða með mælitölum á borð við úrtaksmeðaltal og úrtaksstaðalfrávik (*s*). Því stærra sem úrtakið er því meir mun dreifingin líkjast dreifingu þýðisins sem úrtakið er dregið úr, og mælitölur þess (meðaltal og staðalfrávik) svipa meira til þýðistalnanna.

„*Sampling distribution*“ köllum við úrtakadreifingu á íslensku. Þetta er líkindadreifing einhverrar mælitölu frá einu úrtaki til annars. Yfirleitt er átt við úrtakadreifingu meðaltals (*sampling distribution of the mean*) þótt aðrar úrtakstölur hafi auðvitað einnig úrtakadreifingu.

Úrtakadreifing lýsir breytileika mælitölu milli fjölmargra úrtaka af ákveðinni stærð. Dreifingin gefur til kynna líkindi þess að úrtakstala falli innan ákveðinnar fjarlægðar frá þýðistölunni (t.d. þýðismeðaltalinu). Ef ég veit (eða gef mér) miðsækni úrtakadreifingar meðaltals og breidd dreifingarinnar (staðalvillu meðaltalsins; *standard error of the mean*) get ég ályktað um dreifingu úrtaksmeðaltalanna. Dreifingin í úrtakinu (*sample distribution*) segir mér hins vegar ekkert um meðaltölin heldur aðeins um sjálf stökin (einstaklinganna) í úrtakinu.

Úrtakadreifingin (*sampling distribution*) er því lykilatriði í allri ályktunartölfræði.

Fjallað er ítarlega um úrtakadreifingar í köflum 4.3-4.6 í Agresti (bls. 94–111). Stutt samantekt er á blaðsíðu 107 og samantekt á efni alls kaflans er á bls. 110–111.

1999-03-04 EÖE
1999-03-04 GBA

Hví þarf óháð tilviljunarval í úrtak?

Hvers vegna þarf að velja með tilviljunaraðferð í úrtak og hvers vegna þarf valið að vera óháð?

Við höfum rætt þetta nokkuð ítarlega í tíma, þ.e. að úrtök þurfi að vera dregin með tilviljunaraðferð og með óháðu vali á einstaklingum. Ef þetta er gert, þá vitum við hver úrtakadreifing úrtaksmeðaltals er. Ef þetta er ekki gert, þá hefur úrtakstalan eftir sem áður einhverja úrtakadreifingu en vandinn er sá að við vitum ekki hverjir eiginleikar hennar eru. Við getum því ekki treyst því að niðurstöður útreikninga gefi réttar niðurstöður; strangt til tekið getum við ekki notað hefðbundin marktektarpróf.

Í reynd eru þessar forsendur iðulega brotnar, t.d. byggjast flestar rannsóknir ekki á tilviljunarúrtökum. Þetta leiðir til þess að niðurstöðurnar verða ekki einhlítar, þ.e. töluverður túlkunarvandi fylgir slíkum rannsóknum.

2001-03-23a GBA

Hvað þýðir $p = 0{,}05$?

Í glæru „grunnatriði 18“ er útkoman 5%. Hvað er þessi útkoma að segja mér, þ.e. hvað þýða þessi 5%?

Við tilgátuprófun erum við alltaf að reyna að forðast að gera tvenns konar mistök, (a) að hafna núlltilgátu þegar hún er í raun rétt og (b) að samþykkja núlltilgátu þegar hún er í raun röng. Þegar við tölum um líkur í tilgátuprófum erum við í raun að meta hverjar líkurnar eru á að gera þessi mistök. Fyrri villan er nefnd höfnunarmistök (og er stjórnað með alfamörkunum) en sú seinni er fastheldnimistök. Tíðni höfnunarmistaka stjórnum við með því að velja nægilega lágt α en fastheldnimistökum getum við ekki stjórnað beint.

Ákjósanlegast væri að nota sem lægst α og minnka þannig tíðni höfnunarmistaka. Gallinn er sá að þá eykst tíðni fastheldnimistaka að öðru jöfnu. Yfirleitt er miðað við að α sé annaðhvort 0,05 (5%) eða 0,01 (1%). Það væri best að velja sem lægst α en þar sem það eykur tíðni fastheldnimistaka verðum við yfirleitt að fara einhvern milliveg.

Alfamörkin tengjast *p*-gildinu þannig að ef p er lægra en α getum við hafnað núlltilgátunni. Þannig gefur p líkurnar á því að fá niðurstöður sem víkja frá núlltilgátunni jafnmikið eða meira en meðaltalið sem fékkst í úrtakinu. Það þýðir að ef niðurstaða tölfræðiprófs gefur $p = 0{,}05$ merkir það að 5 prósent líkur séu á þetta miklu (eða meira) fráviki frá núlltilgátunni ef hún er í raun rétt. Á glæru 18 í grunnatriðunum merkir það að miðað við gefin gögn er niðurstaðan í andstöðu við núlltilgátuna; ólíklegt (minna en 5% líkindi) væri að við hefðum fengið þessi gögn við rannsókn ef núlltilgátan væri rétt.

1999-02-26 MDR
1999-02-26 GBA

Þetta þýðir sem sé að það séu 5% líkindi á því að kynjaskiptingin sé jöfn miðað við glæru 18 í Grunnatriðum ályktunartölfræði, eða hvað?

Þetta er algengur misskilningur og er ekki rétt.

Í ályktunartölfræði er sjaldnast sem ég veit líkindi þess að núlltilgátan sé rétt. Hvorki *p*-gildið né α gefur mér tilefni til slíkrar ályktunar. Setjum sem svo að p sé 0,05. Þá veit ég að það eru aðeins 5% líkindi á að fá niðurstöðuna (eða meira frávik frá núlltilgátunni) í þeim tilvikum þar sem núlltilgátan er rétt. Hversu líkleg núlltilgátan er eða yfirhöfuð hvort hún er rétt eða röng er eitthvað sem ég ekki veit og get ekki lesið út úr niðurstöðunum.

1999-02-26 GBA

Fyrir hvað stendur π?

Í umræðu um öryggisbil fyrir hvað stendur π (pí)?

Pí stendur fyrir þýðishlutfallið. Athugaðu að grísku stafirnir standa yfirleitt fyrir þýðistölur; úrtakstölur eru táknaðar með samsvarandi rómverskum stöfum (í þessu tilviki er úrtaksstaðalfrávikið táknað með p).

(2002-03-11b GBA)
2003-01-11e GBA

Öryggisbil þýðis- og úrtakstalna

Ég er að gera eitthvað vitlaust þegar ég geri öryggisbil þýðishlutfalls. Hver er munurinn á π og π með haka fyrir ofan? Hver er munurinn á X með striki yfir og μ? Þetta á við um dæmi 21 og 44.

Hvað er öðruvísi við útreikninginn miðað við þýði heldur en þegar miðað er við úrtak?

Grísku stafirnir vísa til þýðistalna (*parameter*) en rómönsku stafirnir til úrtakstalna; grískir stafir með hatti (haka) standa fyrir spátölur (*estimate*).

Þýðishlutfallið er því táknað með gríska stafnum π, úrtakshlutfallið með p og spátalan fyrir þýðishlutfallið er táknað með $\hat{\pi}$. Besta spágildið fæst að jafnaði með því að nota úrtakshlutfallið; úrtakshlutfallið og spátalan fyrir þýðishlutfallið er því ein og sama talan og því er $\hat{\pi}$ oft notað í staðinn fyrir p.

Á sama hátt er úrtaksmeðaltalið táknað með $\overline{X}$, þýðismeðaltalið með μ og spágildi þýðismeðaltalsins með $\hat{\mu}$; úrtaksmeðaltalið er að jafnaði besta spáin fyrir þýðismeðaltalið og því er $\overline{X}$ og $\hat{\mu}$ gjarnan sama talan.

Öryggisbil fyrir þýðistölu er reiknað á svipaðan hátt og öryggisbil fyrir úrtakstölu. Í báðum tilvikum reiknum við staðalvillu til að meta breidd úrtakadreifingarinnar. Yfirleitt vitum við staðalfrávik þýðis þegar öryggisbil úrtakstölu (t.d. úrtaksmeðaltals) er reiknað og því er miðað við *z*-dreifingu. Hins vegar vitum við að jafnaði ekki um þýðisstaðalfrávikið þegar við reiknum öryggisbil þýðistölu (t.d. þýðismeðaltals) og miðum því við *t*-dreifingu í útreikningunum. Að vísu er töluverð hefð fyrir því að nota alltaf *z* þegar reiknað er öryggisbil þýðishlutfalls, enda eru úrtökin yfirleitt það stór að munurinn á *t* og *z* er orðinn lítill sem enginn.

Að öðru leyti er munurinn fólginn í staðsetningu öryggisbilsins. Þegar við reiknum öryggisbil úrtakstölu þekkjum við þýðistöluna en viljum meta hver sé líkleg dreifing samsvarandi úrtakstölu frá einu úrtaki til annars. Öryggisbil úrtakstölu er alltaf staðsett með miðjuna þar sem þýðistalan er; öryggisbil úrtaksmeðaltals er miðjað á þýðismeðaltalið og öryggisbil úrtakshlutfalls er miðjað á þýðishlutfallinu.

Þegar við reiknum öryggisbil þýðistölu þekkjum við að jafnaði úrtakstölurnar en viljum ákvarða það talnabil sem er líklegt til að innihalda (fanga) þýðistöluna. Breidd úrtakadreifingarinnar segir til um hversu breitt talnabilið ætti að vera en við þurfum einnig að staðsetja það þannig að hámarkslíkur séu á því að þýðistalan lendi innan þess. Þar sem samsvarandi úrtakstala gefur að jafnaði bestu spá um þýðistöluna, næst þetta með því að miðja bilið á úrtakstöluna. Því miðjum við öryggisbil fyrir þýðismeðaltal á úrtaksmeðaltalinu og öryggisbil fyrir þýðishlutfall á úrtakshlutfallinu.

Það er margt líkt með þessum tveimur tegundum öryggisbila en einnig ýmislegt sem er ólíkt. Því getur verið auðvelt að ruglast í útreikningum og túlkun.

2001-04-10a GBA

Hvort er betra 95% eða 99% öryggisbil?

Hvernig veit ég hvort er betra að hafa 95% eða 99% öryggi öryggisbil?

Þegar öryggisbil er reiknað þarf að velja það öryggi sem miða skal við. Augljóslega er betra að hafa 99% heldur en 95% öryggi. Aukið öryggi þýðir hins vegar einnig að bilið stækkar; öryggisbilið er því mun víðara fyrir 99% öryggi heldur en það er fyrir 95% öryggi. Það er því lítil stoð í miklu öryggi ef það er samfara mjög víðu bili.

Öryggisbil fela í sér eitthvert talnabil sem ég vonast til að muni innihalda þýðistöluna, t.d. þýðismeðaltalið. Því þrengra sem bilið er því gagnlegra er það þar sem það felur í sér að ég staðset þýðistöluna nákvæmar en þegar bilið er vítt. Þröngt bil er þó því aðeins gagnlegt ef ég hef einhverja vissu fyrir því að bilið innihaldi þýðistöluna í reynd. Þröngt bil samfara mikilli óvissu er því ekki gagnlegt.

Við höfum því val á því að auka öryggið og víkka með því bilið eða að minnka öryggið og þrengja þannig bilið. Þetta þarf að vega hvort á móti öðru þannig að gagnleg niðurstaða fáist. Ef staðalvillan er há, myndi ég hneigjast til þess að hafa minna öryggi til að vinna á móti þeirri víkkun bilsins sem leiðir af staðalvillunni. Ef staðalvillan er lág, get ég hins vegar hækkað öryggið án þess að það verði til þess að öryggisbilið verði of vítt.

Þannig fer það eftir aðstæðum og markmiðum mínum hverju sinni hvort ég ætti að velja 95% eða 99% öryggisbil. Ég get litið svo á að vídd bilsins feli í sér hversu nákvæmlega ég staðset þýðistöluna en öryggið felur í sér óvissu um það hvort bilið innihaldi þýðistöluna í reynd. Ég þarf því að vega nákvæmnina á móti óvissunni. Hvernig það skal gert í hverju og einu tilviki hefur ekkert eitt svar.

2001-04-10c GBA

Er hægt að hafa 51,3% af 100 einstaklingum?

Í einu dæmi í Stoðkverinu stendur að það hafi verið tekið hundrað manna úrtak. Í ljós kom að 51,3% nemenda skrifuðu „dalsminni“ í stað „dalsmynni.“ Hvernig geta 51,3% af 100 nemendum skrifað þetta, þ.e. 51,3 nemandi (í stað 51 nemenda)?

Er þetta mistök eða vitlaust skilið hjá mér?

Þetta er rétt skilið. Auðvitað er þetta óhugsandi. Þetta er glappaskot en þér er óhætt að horfa alfarið fram hjá því. Einfaldlega reiknaðu dæmið eins og þetta sé í raun 51,3 einstaklingur jafnvel þótt það sé í reynd óhugsandi að skipta einum nemanda svona í þrennt.

2003-03-21c GBA

Dæmi 5.7 í Agresti

Ég skil ekki alveg b-lið í dæmi 5.7 í Agresti. Hvernig fær hann út að fjórfalda þurfi úrtaksstærð til að fá helmingi þrengra öryggisbil? Ég sé ekki hvernig upplýsingarnar eiga að duga til að geta reiknað þetta út.

Staðalvillan er fengin með því að deila kvaðratrótinni af úrtaksstærðinni í staðalfrávikið. Staðalvillan minnkar því í öfugu hlutfalli við kvaðratrótina af úrtaksstærðinni. Ef úrtakið fjórfaldast þá minnkar staðalvillan því í hlutfallinu $1/\sqrt{4}$. Kvaðratrótin af 4 er 2 og því helmingast staðalvillan við það að fjórfalda úrtaksstærðina. Ef staðalvillan helmingast, þá verður öryggisbilið einnig helmingi þrengra.

Skoðaðu nokkur öryggisbilsdæmi og sjáðu sjálf hvað gerist þegar þú stækkar eða minnkar úrtakið. Ef þú tvöfaldar úrtakið, ætti staðalvillan að minnka í hlutfallinu $1/\sqrt{2}$= 0,707, þ.e. minnka um tæplega 30%. Ef þú helmingar úrtakið ætti staðalvillan að breytast um hlutfallið $1/\sqrt{½}$= 1,41, þ.e. stækka um 41%. Öryggisbilin ættu að breytast í sömu hlutföllum. Prófaðu raunveruleg dæmi og sjáðu hvað gerist ef þú reiknar þau til enda með mismunandi úrtaksstærðum. Það er góð æfing í því að skilja til fulls hvernig þetta virkar. Sá skilningur skilar sér seinna, bæði á prófi í Aðferðafræði II en einnig þegar þú ferð að skipuleggja eigin rannsóknir eða túlka rannsóknarniðurstöður annarra.

2001-03-08a GBA

Í a lið er gefið N sem er jafnt og 100, og meðaltal úrtaks, sem er jafnt og 5,3. Síðan er á það bent, að með endurteknum úrtökum, þá ætti meðaltal úrtaksins ekki að fara lengra frá raungildinu, nema sem nemur 1 í 95% tilvika, þ.e. 4,3–6,3.

Hvað þarf að reikna og hvernig er best að nálgast áframhaldið? Gott væri á fá ábendingar.

Í stofni spurningarinnar kemur fram túlkun á öryggisbilinu sem myndað var, þ.e. að úrtaksmeðaltalið hafi verið 5,3 og að búast við að úrtaksmeðaltalið myndi falla innan einnar einingar frá þýðismeðaltalinu í 95% tilvika ef dregin væru mörg tilviljunarúrtök af þessari stærð.

Úrlausnarefni spurningarinnar er einfaldlega að skilja þetta orðalag og sýna skilninginn með því að setja fram og túlka öryggisbilið sem lýst er með þessum orðum. Ég veit ekki hvort það hjálpar, en veltu því fyrir þér hvað öryggisbilið þurfi að vera breytt til að úrtaksmeðaltalið falli í 95% tilvika innan einnar einingar (eins dags) frá þýðismeðaltalinu. Úrlausnarefnið snýst sem sé um túlkun öryggisbila.

2001-03-13a GBA

Hvers vegna er staðalfrávik ekki gefið upp í dæmi 43?

Ég er búin að vera að undirbúa mig fyrir hlutapróf II og er búin að lesa glærurnar og alltaf þegar reikna á öryggisbil þar er gefið meðaltal og staðalfrávik. Nú þegar ég er að reikna í Stoðkverinu þá eru þessar mælitölur aldrei gefnar og ég finn ekki almennilega hvernig ég á að reikna þær.

Eru þessar mælitölur aldrei gefnar á prófum eða er þetta bara svona í Stoðkverinu? Ég er að reikna dæmi 43 í Stoðkverinu og finn ekki hvernig ég á að finna staðalfrávik og staðalvillu ef ég er bara með N= 650 og hvernig þessi 650 skiptast í 6 flokka.

Í dæmi 43 færðu uppgefið fylgi hvers stjórnmálaflokks í prósentum. Þú hefur því úrtakshlutfall fyrir hvern flokk og notar formúlur fyrir staðalfrávik og staðalvillu hlutfalla til að reikna öryggisbil fyrir hvern flokk fyrir sig.

2001-03-17a GBA

Hvernig reikna ég dæmi 43 í Stoðkveri?

Geturðu sagt mér hvernig ég get reiknað dæmi 43 í Stoðkverinu? Ég veit að ég á að nota staðalfrávik og staðalvillu hlutfalla, en ég finn hvergi formúlu fyrir staðalfrávik hlutfalla! Einhvers staðar í glærunum er talað um að hægt sé að reikna beint staðalvillu, en það er í sambandi við tvíkostabreytur sem á ekki við í þessu dæmi.

Formúlur finnurðu í liðum e, f, g og h undir fyrirsögninni **Öryggisbil** í **Formúlublaði í öðrum þriðjungi**. Við gerum ráð fyrir að þetta sé túlkað sem röð tvíkostabreyta. Það eru 14% sem ætla að kjósa A-lista og því 86% sem ætla það ekki; 18% ætla að kjósa B og því 82% sem ætla það ekki; o.s.frv.

Það er rétt hugsað hjá þér að þetta séu í reynd margkostabreyta. Strangt til tekið ættum við því ekki að meðhöndla þetta sem röð tvíkostabreyta. Í reynd er það oft gert, t.d. í öllum skoðanakönnunum. Þetta verklag hefur ekki afgerandi áhrif á niðurstöður og því er þér óhætt að reikna dæmið eins og það er lagt upp hér.

2003-03-19a GBA
2003-12-29e GBA

Hvernig túlka ég öryggisbil?

Hvernig túlka ég öryggisbil eins og í dæmi 52?

Dæmigerð túlkun er gefin upp í svörum fyrir dæmi 50.

2001-04-10d GBA

Túlkun öryggisbils fyrir úrtaksmeðaltal

Er eftirfarandi rétt túlkun fyrir úrtaksmeðaltal: 95% líkur á því að úrtaksmeðaltalið sé á bilinu $A < \overline{X} < B$?

Já, mér sýnist þetta vera rétt.

Vor 1999 GBA

Er eftirfarandi rétt túlkun fyrir úrtaksmeðatal? Að jafnaði munu 95% öryggisbila byggð á sams konar óháðum úrtökum innihalda meðaltal þýðisins.

Nei, þetta ætti við um öryggisbil *þýðis*meðaltals.

Vor 1999 GBA

Túlkar Agresti öryggisbil rétt?

Í dæmi 5.2 bls. 133 í Agresti er reiknað út öryggisbil fyrir hlutfall þeirra sem eru samþykkir fóstureyðingum. Samkvæmt því er hægt að fullyrða með 95% öryggi að 44%–48% fólks finnist fóstureyðingar réttlætanlegar. Svona myndi ég allavega túlka niðurstöðuna en Agresti segir að hlutfall þeirra sem samþykkja fóstureyðingar í þýði sé a.m.k 44% og í mesta lagi 48%. Hvernig getur hann fullyrt þetta? Miðað við 95% öryggi (en ekki 100) hljóta að vera smá líkur á því að öryggisbilið innihaldi ekki þýðishlutfallið og það því minna en 44% eða stærra en 48%.

Það væri gott ef þú gætir útskýrt hvers vegna Agresti má taka svona til orða.

Mér sýnist Agresti einfaldlega beygja út í skurð þarna. Það virðist einfaldast að skilja orðalag Agrestis þannig að þýðishlutfallið geti ekki verið minna en 44% né hærra en 48%. Hvoru tveggja er beinlínis rangt eins og þú réttilega bendir á.

Þú ert að skilja þetta efni rétt; textinn í námsbókinni er rangur.

16-03-2001a GBA

‡Öryggisbil fyrir heimilisofbeldi

Ég átta mig ekki á tölunum á glærunni Athugun á heimilisofbeldi í Grunnatriði í ályktunartölfræði. Þar er sagt að á bilinu 986 til 2.285 manns hafi orðið fyrir heimilisofbeldi af hendi maka á síðustu 12 mánuðum. Punktspáin er sögð vera 1.750.

Það sem vefst fyrir mér er að punktspáin er ekki í miðju öryggisbilinu; miðja bilsins væri 1.636. Hvernig stendur á þessu?

Þær formúlur sem þið fáið í Aðferðafræði II miðast alltaf við samhverf öryggisbil. Það er alls ekki sjálfgefið að öryggisbilin séu samhverf, sérstaklega þegar unnið er með lág hlutföll.

Almennt séð verðum við að telja að það sé erfiðara að fara frá t.d. 1% niður í 0,5% heldur en upp í 2%. Þú getur hugsað það þannig að þú sért með verksmiðju þar sem eitt stykki af hverjum hundrað sé gallað. Það væri mjög erfitt að bæta framleiðsluna, lækka hlutfall gallaðra framleiðsluvara. Hins vegar þyrfti ekki nema að sofna aðeins á verðinum og þá gæti hlutfall gallaðra framleiðsluvara farið upp í 2% eða jafnvel enn hærra.

Öryggisbilin og staðalvillurnar sem koma út úr formúlunum í Aðferðafræði II taka ekki tillit til þessa eiginleika lágra hlutfalla (og samsvarandi eiginleika hárra hlutfalla). Reglan að lægra hlutfallið þurfi að samsvara a.m.k. 10 einstaklingum (stökum) kemur þó að miklu leyti á móts við þetta, þ.e. reynir að tryggja að við munum aðeins nota samhverf öryggisbil þegar það á við. Reglan er þó aðeins viðmið og stundum ættu öryggisbilin að vera ósamhverf þótt 10 eða fleiri einstaklingar séu í smærri hópnum.

Nákvæmustu útreikningar sem ég kemst í benda til þess að öryggisbilið eigi að vera 0,00657 – 0,01523 eða frá 0,7% upp í 1,5%. Það er því prentvilla á glærunni en hún segir ranglega að efri mörkin séu 1,3%. Þetta skýrir það að fjöldatölur eru ósamhverfar.

Á hinn bóginn sé ég að allar fjöldatölur eru á reiki. Það er því líkast að ég hafi ýmist miðað við 150.000 fullorðna Íslendinga eða 175.000. Ég vil þó ekki breyta glærunni fyrr en ég kemst í frumheimildirnar. [Ástæða þessa er viss ónákvæmni í skýrslunni sem upplýsingarnar eru komnar úr. Gefnar eru upp fjöldatölurnar 1.000–1.100 konur og 650 karlar. Hins vegar er ekki rökstutt hvernig þær tölur eru fengnar. Þetta veldur ósamræmi. 2001-.02-23 GBA]

Allt eru þetta þó aukaatriði miðað við aðalatriðið sem glærunni og umfjöllun um hana er ætlað að benda á. Við meðhöndlum oft niðurstöður rannsókna eins og fengin sé endanleg óbreytanleg niðurstaða. Í reynd er mikil óvissa um niðurstöðurnar og eðlilegra að líta svo á að þær hafi verið staðsettar á ákveðnu talnabili fremur en að ein tala hafi verið fastákvörðuð. Rannsóknin sem glæran byggir á er sennilega eitt skýrasta dæmið um slíkt ofmat á nákvæmni niðurstaðna.

2000-03-22 GBA

Öryggisbilið hefur verið leiðrétt á glærunni.

Ónákvæmni í fjöldatölum á glærunni er tilkomin vegna ófullnægjandi upplýsinga í skýrslunni sem hún byggir á. Gefnar eru upp fjöldatölurnar 1.000–1.100 konur og 650 karlar. Hins vegar er ekki sýnt nákvæmlega hvernig þær tölur eru fengnar. Þetta veldur ósamræmi milli útreiknaðs öryggisbils og þeirra fjöldatalna sem birtast á glærunni. Ósamræmið hefur þó engin áhrif á megininntak glærunnar og því einfaldast að horfa framhjá því.

2001-02-23 GBA

Furðuleg svör við dæmi 44 í Stoðkveri?

Mér finnast svörin við dæmi 44 í Stoðkverinu dálítið furðuleg. Í dæminu kemur t.d. ekki fram hvort úrtakið sé tilviljanaúrtak en samt á að reikna öryggisbil. Þetta ruglar mig því mér finnst að þetta ætti að koma fram. Ef þetta dæmi kæmi á prófi myndi ég halda að þetta væri svona „trick question.“

Að auki er brottfall úr úrtakinu en niðurstöðurnar eru einungis miðaðar við þá sem eru eftir. Ekkert er vitað um þá sem féllu burt svo er tæknilega séð nokkuð hægt að fullyrða að með 95% öryggi sé þýðishlutfallið á uppgefnu bili (eins og sagt er í niðurstöðum)? Þeir sem féllu út gætu verið allt öðruvísi en hinir (t.d. mjög á móti rafmeðferð og þess vegna ekki viljað taka þátt).

Athugaðu það vel að það eru engar gildrur í prófum í Aðferðafræði II. Það er miðað við að allar spurningar eigi sér eðlileg svör og þeir sem eru vel að sér sjái nokkuð glögglega hvert hið eðlilega svar er. Auðvitað getur okkur verið mislagðar hendur og spurningar verið gallaðar eða jafnvel misheppnast, en það getur aldrei verið markmið kennara að búa til spurningar sem villa um fyrir nemendum. Slíkt væri einfaldlega tilraun til að skemma fyrir sér og markmiðum starfsins.

Ályktunartölfræði byggir á tilviljunarúrtökum. Sú forsenda er hins vegar nær alltaf brotin í reynd. Það leiðir til túlkunarvanda. Í Aðferðafræði II er þér hins vegar almennt óhætt að gefa þér að um tilviljunarúrtak sé að ræða. Við getum ekki þulið upp allar forsendur við hvert og eitt dæmi. Við reynum hins vegar láta þessa forsendu koma fram eins oft og við komum því við. Ef þetta veldur þér vandræðum, sting ég upp á að þú setjir forsenduna í svarið, þ.e. segir sem svo: „Ef gera má ráð fyrir að þetta sé tilviljunarúrtak ….“

Þetta með brottfallið er hárrétt athugað. Oftast er gert ráð fyrir því að brottfallið sé tilviljunarkennt, en það er þó fjarri því víst að svo sé. Það eru til mikil fræði um brottfall og ólíkar forsendur varðandi eðli þeirra. Við getum hins vegar yfirleitt skýlt okkur bak við það að brottfallið sé annað hvort tilviljunarkennt eða ef svo er ekki að þá sé alhæft yfir á annað þýði. Þegar við leggjum t.d. fyrir spurningalista erum við í reynd að alhæfa yfir á þá sem svara slíkum lista. Það þýði er ólíkt hinu; ef það er mjög ólíkt þá hefur það alvarlegar afleiðingar fyrir túlkun niðurstaðnanna.

Viðlíka vandi við túlkun niðurstaðna er alltaf fyrir hendi í öllum rannsóknum en í mismiklum mæli. Við getum ekki forðast hann en okkur ber að lágmarka hann. Við þurfum að vita af honum og hafa hann í huga við túlkun niðurstaðnanna. Í svörum við efnisspurningum í Aðferðafræði II getur hins vegar annað hvort litið fram hjá honum eða fléttað hann inn í svarið, t.d.: „Ef gera má ráð fyrir að brottfallið sé tilviljunarkennt ….“

2003-03-21b GBA
2003-12-29c GBA

‡Er þetta rétt túlkun á öryggisbili?

Ég er með spurningu um túlkun öryggisbila. Mig langar að taka smá dæmi. Segjum sem svo að ég hafi reiknað 95% öryggisbil fyrir þýðismeðaltal og fengið út að það sé á bilinu X–Y. Væri þá rétt að segja:

1. Með 95% öryggi er hægt að segja að þýðismeðaltalið lendi á þessu bili.
2. Ef úrtök væru endurtekið dregin úr sama þýði myndu 95% af reiknuðum öryggisbilum úr þeim innihalda hið rétta þýðismeðaltal.
3. Líkurnar eru því aðeins 1/20 að það öryggisbil sem var reiknað geri það ekki.

Þetta sem ég sagði síðast finnst mér leiða af þessum fyrri fullyrðingum mínum, en ef ég hef ekki misskilið þig þá sagðir þú í tíma að þetta væri ekki rétt. Mig langar að vita af hverju …

Fyrstu tveir liðirnir eru réttir en ekki sá þriðji. Ástæðan er dálítið flókin og vandútskýrð. Leyfðu mér samt að reyna það.

Tölfræði miðast við að litið sé á líkindi sem hlutfallslega tíðni, þ.e. hversu oft atburður gerist að jafnaði sem hlutfall af heildafjölda þeirra atburða sem eru til athugunar. Þetta er ólíkt huglægum líkindum: Ef ég segi að það séu yfirgnæfandi líkur á að Bandaríkin sigri Írak, byggist það á einhvers konar fullvissu minni en ekki hlutfallslegri tíðni. Það einfaldlega er engin endurtekning og því engin hlutfallsleg tíðni. Við getum talað um sennileika til að aðgreina huglægu líkindin frá hlutfallslegri tíðni. Það hafa staðið yfir deilur um þessi hugtök áratugum saman innan tölfræðinnar og gildi þeirra hvors fyrir sig. Sú aðferðafræði sem félagsvísindin byggja á miða hins vegar skýrlega við hlutfallslega tíðni sem grunninn fyrir líkindum.

Ef ég dreg endurtekin úrtök úr sama þýði, þá hefur það skýra merkingu að öryggisbil innihaldi annað hvort þýðistöluna eða ekki. Það hefur því skýra merkingu að tala um að 95% öryggisbilanna innihaldi þýðistöluna. Hér erum við því með skýrt dæmi um hlutfallslega tíðni. Liður 2 í spurningunni er því réttur og hefur skýra merkingu út frá hlutfallslegri tíðni.

Ef við tökum eitt tiltekið öryggisbil, vandast málið. Annað hvort inniheldur það þýðistöluna eða ekki. Ef það inniheldur ekki þýðistöluna, er ekki hægt að segja að samt séu 95% líkindi á að þýðistalan sé innan bilsins. Kannski er einfaldast að hugsa þetta út frá dæmi.

Ég er karl. Væri hægt að segja að það væru um 50% líkindi á því að ég væri kona? Auðvitað er hugsanlegt að ég villi á mér heimildir og sé kona þrátt fyrir ytri vísbendingar um karlkyn; það eru einhver dæmi um slíkt. En væri slíkt ekki dæmi um að ég væri 100% kona sem léti eins og hún væri karl. Það hljóta því að vera annað hvort 100% líkur á því að ég sé karl eða 100% líkur á því að ég sé kona. Athugaðu að hér er ég að nota hugtakið í merkingunni hlutfallsleg tíðni.

Áður en ég fæddist hafa væntanlega verið um 50% líkur á að ég væri kona. Miðað við að ég var ófæddur og kynið ekki þekkt, var hægt að miða við það að þegar til lengdar er litið þá eru tæp 50% nýbura stúlkur. Þar til ég fæddist gat ég því verið hvort kynið sem var og því var hlutfallsleg tíðni atburðarins að barnið reyndist vera sveinbarn um 50%.

Snúum nú málinu aftur að öryggisbilum. Áður en úrtakið er tekið og öryggisbil reiknað, getur það annað hvort innihaldið þýðistöluna eða ekki. Í þeim skilningi eru því 95% líkur á því að það innihaldi þýðistöluna. Eftir að tiltekið bil hefur verið reiknað, þá er ekki lengur um líkindi að ræða fremur en hægt er að ræða líkur þess að ég sé annað en það sem ég er.

Ef við ræðum öryggisbil fyrir þýðismeðaltal sérstaklega, fær þessi umræða aukna dýpt. Í endurteknum úrtökum verða sum bilin mjög víð en önnur afar þröng, sum eru nokkurn veginn rétt staðsett en önnur allt of ofarlega eða neðarlega á talnalínunni. Setjum nú sem svo að ég taki úrtak 20 karlmanna og reikni meðaltal fyrir líkamshæð og samsvarandi öryggisbil. Ég gæti lent á þröngu öryggisbili, segjum 179–181. Ég gæti einnig lent á víðu bili, t.d. 170–190. Ef við hugleiðum þessi tilteknu bil, hvernig má það vera að það séu jafnmiklar líkur á því að þýðismeðaltalið sé á þrönga og á víða bilinu? Ef það eru 95% líkur á því að líkamshæð karla sé 179–181 cm, hljóta þá ekki að vera enn meiri líkur á að bilið 170–190 innihaldi þýðismeðaltalið? Athugaðu að þótt öryggisbilin séu þetta misvíð, eru tvö alveg sams konar 95% öryggisbil en reiknuð í tveim ólíkum úrtökum. Bersýnilega er eitthvað bogið við það að tala um líkindi þegar miðað er við ákveðin tiltekin talnabil.

Þennan vanda er hægt að leysa formlega en ekki ef miðað er við hlutfallslega tíðni. Lausnin felst í því að hugsa um öryggisbil sem sennileikabil, þ.e. að það séu ákveðin huglægur sennileiki á því að bilið innihaldi þýðistöluna. Það yrði of flókið að skýra þau fræði nákvæmlega. Látum nægja að ef þú talar um að tiltekið öryggisbil innihaldi þýðismeðaltal með 95% öryggi eða að þetta sé sennileikabil, þá er hægt að verja þá orðanotkun. Ef þú hins vegar vilt tala um líkindi í sambandi við tiltekin bil, þá ertu komin í vandræði og lendir í óþægilegum mótsögnum.

2003-03-21a GBA
2003-12-29d GBA

Nauðsynleg úrtaksstærð

Varðandi úrtaksstærð þá skil ég einfaldlega ekki glæru 24 [Nauðsynleg úrtaksstærð] í öryggisbilum þar sem þú talar um hlutföllin, hver er nauðsynleg úrtaksstærð?

Já, kannski er þetta eitthvað ruglingslegt. Ég átta mig þó ekki á því hvernig ég gæti útskýrt þetta nánar. Glæra 25 [Útreikningur úrtaksstærðar] ætti að setja þetta í rétt samhengi.

Í aðalatriðum byggist þetta á því að ég er að skipuleggja rannsókn og vil að niðurstöðurnar hafi ákveðna lágmarksnámkvæmni. Nákvæmnin ræðst aðallega af úrtaksstærðinni. Því vil ég geta reiknað út hversu stórt úrtak ég þarf að velja til að fá þá nákvæmni sem ég vil fá.

Nákvæmnin felst í líklegri stærð öryggisbilsins; ég vil vita líklega stærð þess *áður* en ég framkvæmi rannsóknina því ég vil að úrtakið sé nægilega stórt en þó ekki stærra en þörf krefur.

Vor 2000 GBA

Túlkun á marktekt í tölvuverkefnum

Ég var að lesa aftur yfir SPSS verkefnin, því þau gefa oft raunverulega mynd. Þar stoppaði ég á verkefni 4, sem spurt er hvaða undirpróf virðast hafa annað þýðismeðaltal en erlenda viðmiðsgildið.

Svar mitt er að þau atriði sem sýna 0,000 og 0,017 skeri sig marktækt frá erlenda viðmiðinu.

Já, þetta hljómar sannfærandi. En ekki segja að þau skeri sig marktækt frá viðmiðinu; betra væri að segja að þú ályktir að þú séu ólík viðmiðsgildinu í þýði vegna þess að *t*-prófið er marktækt. Þessi ofurnákvæmni tengist næsta skrefi í spurningu þinni.

2000-04-29 GBA

Þýðir þetta að núlltilgátunni er hafnað varðandi þessi atriði og að munur sé á frammistöðu íslenskra barna og þeirra erlendu? Það sem truflar mig er að fyrir ofan stendur „líkur á niðurstöðu ef núlltilgátan er rétt.“

Einmitt. Þú kemst að því að frávik niðurstöðunnar frá viðmiðsgildinu er ólíklegt ef núlltilgátan er rétt. Það gefur þér *leyfi* til að staðhæfa að núlltilgátan sé röng. En þú *veist* ekki hvort hún er röng eða rétt.

Auk þess hefur þú í sjálfu sér engan áhuga á niðurstöðunni í úrtakinu; t.d. skiptir engu máli hvort þau „skeri sig marktækt frá.“ Marktekt er eitthvað sem gerist í þessu tiltekna úrtaki, en þú leitar upplýsinga um þýðið. Þitt vilt vita hvort þér sé stætt á því að fullyrða að núlltilgátan sé rétt. Þér er heimilt að hafna henni þegar t-prófið er marktækt.

Í þessum skilningi er marktektarpróf aðeins tæki sem þú notar til að draga ályktanir um þýðið. Marktektin sem slík er einskis virði ef gerir þér ekki kleift (leyfir þér) að álykta (réttilega eða ranglega) um þýðið.

2000-04-29 GBA

Alfa og beta

Ég er að reyna að átta mig á Aðferðafræðinni. Það er svo auðvelt að rugla þessu fram og til baka. Ég er að reyna að böggla saman punktum úr fyrirlestrinum **Ályktanir í einum hópi**. Eru eftirfarandi fullyrðingar t.d. réttar hjá mér? Ef þessar fullyrðingar eru réttar er ég aðeins að skilja þetta.

Ég vil fá lágt β en hátt 1–β?

Já!

Með því að stækka úrtakið þrengi ég kúrfuna og lækka β en hækka 1–β?

Já!

Því hærra sem α er því stærra verður β og minni afköst þ.e. minni líkur á því að ég hafni núlltilgátunni ef hún er röng?

Nei, það er öfugt. Því *lægra* alfa (t.d. 0,01 í stað 0,05) því stærra beta og því minni líkur á að hafna núlltilgátunni ef hún er röng (þ.e. lægra 1–β).

Ég vil fá lágt α (alfa) en hátt 1–β?

Já, það er rétt. Því miður lækkar 1–β við það að α lækkar.

2002-03-11a GBA

Hvenær er *t*-dreifing notuð?

Notar maður t-dreifingu bara þegar þýðistölur (staðalfrávik þýðismeðaltals) eru óþekktar?

Það hefur nefnilega vafist fyrir okkur að reikna sum dæmin í Agresti því að það virðist sem z-dreifingin sé notuð þótt úrtakið sé minna en 30. Það sem við héldum var að úrtak sem væri minna en 30 væri aðalskilyrðið fyrir notkun t-dreifingar.

Sjá t.d. dæmi varðandi öryggisbil í kafla 5; það er ekki fyrr en í kafla 6 sem talað er um t-dreifingu.

t-dreifing er aðeins notuð þegar staðalfrávik þýðis eru óþekkt, það er við ákvörðun öryggisbils fyrir þýðismeðaltal og við tilgátuprófun. Agresti fjallar um t-dreifingu í 6. kafla þar sem hann fjallar meðal annars um hvernig á að finna öryggisbil þýðismeðaltals með litlu úrtaki ($N < 30$).

Í öllum dæmum í fimmta kafla þar sem $N < 30$ og Agresti beitir z-dreifingu, er hann væntanlega að reikna út öryggisbil úrtaksmeðaltala, en þá er staðalfrávik þýðis þekkt og réttilega stuðst við z-dreifingu (normaldreifingu).

Ef úrtaksstærðin er nægjanleg, t.d. 30 einstaklingar eða meira, þá gefur t-dreifing sömu niðurstöður og z-dreifing (normaldreifing). Rökréttast er að nota samt t ef staðalfrávik þýðis er óþekkt jafnvel þótt úrtakið sé þetta stórt. Agresti velur hins vegar að nota alltaf z-próf í stórum úrtökum hvort sem staðalfrávik þýðis er þekkt eða ekki; það kemur ekki að sök og gefur sömu niðurstöður svo fremi sem úrtaksstærðin sé 30 eða meira.

1999-03-05 EÖE
1999-03-05 GBA

Hvernig getur *t*-dreifing byggst á normaldreifingu?

Er það rétt að það er t-dreifing í t-prófi en krafan er samt þegar ég ætla að nota það próf að það sé normaldreifing í þýði eða úrtak nægilega stórt til að tryggja normaldreifingu?? Er þetta algert rugl? Kæmi mér ekki á óvart!

Já, þetta er rétt. t-dreifing lýsir meðaltölum úrtaka sem fengin eru úr normaldreifðu þýði. Þetta kann að hljóma undarlega en ef þú hugsar málið sérðu að það er ekki rugl. Normaldreifingin á við um þýðið, t-dreifingin á við um meðaltöl úrtaka dregnum úr þessu þýði. Hér er því ekki um raunverulega mótsögn að ræða.

Ef við þekkjum gildi þýðisstaðalfráviksins þá getum við gert ráð fyrir normaldreifingu á úrtakadreifingunni; ef þýðisstaðalfrávikið er óþekkt verðum við að ganga út frá t-dreifingu. Dreifing meðaltalanna er auðvitað alltaf ein og sú sama en vegna skorts á upplýsingum verðum við að ganga út frá víðari úrtakadreifingu en ella þegar okkur skortir upplýsingar um þýðisstaðalfrávikið.

2001-04-03c GBA

Lögun *t*-dreifingarinnar

Hvers vegna er hali t-dreifingar lengri en normaldreifingar?

Fjallað er um þetta efni í kafla 6.5 í Agresti. Athugaðu sérstaklega þriðja ótölusetta liðinn á bls. 181.

Að ákveðnum forsendum gefnum lýsir normaldreifing úrtakadreifingu meðaltala. Breidd úrtakadreifingarinnar fæst með því að reikna staðalvillu meðaltala. Ef þýðisstaðalfrávikið er þekkt fást þannig nákvæmar upplýsingar um breytileika úrtaksmeðaltala dregnum úr viðkomandi þýði. Ef við höfum úrtaksmeðaltal en þekkjum ekki þýðismeðaltalið, gefur breidd úrtakadreifingarinnar okkur því upplýsingar um þá óvissu sem er samfara því að nota úrtaksmeðaltalið sem spátölu fyrir þýðismeðaltalið.

Ef staðalfrávik þýðis er óþekkt, vitum við ekki hver staðalvilla meðaltala er. Við getum hins vegar áætlað það á grundvelli úrtaksstaðalfráviksins, þar sem úrtaks-staðalfrávikið er spátala fyrir þýðisstaðalfrávikið. Sú spá er háð óvissu sem leiðir til þess að við getum aðeins áætlað (spáð fyrir um) staðalvilluna með ákveðinni óvissu. Við vitum því ekki nákvæmlega hver breidd úrtakadreifingarinnar heldur höfum við aðeins óvissa spá um hana.

Þar sem staðalvillan er háð ákveðinni óvissu þarf að gera ráð fyrir henni við allar ályktanir. Það er gert með því að nota *t*-dreifingu í stað normaldreifingar þegar staðalfrávik þýðisins er óþekkt. Óvissan leiðir til þess að spá okkar um meðaltal þýðis verður óvissari sem samsvarar því að við þurfum að gera ráð fyrir því að breidd úrtakadreifingarinnar sé meiri heldur en spágildi staðalvillunnar gefur til kynna. Til að koma á móts við þetta er *t*-dreifingin breiðari en normaldreifing. Eftir því sem úrtakið er minna því meiri óvissa er um staðalvilluna og því breikkar *t*-dreifing eftir því sem frígráðunum fækkar.

2002-05-02b GBA
2003-01-11a GBA

Frávillingar og eiginleikar meðaltala

Þú fjallaðir einnig um meðaltöl á einkunnum sex úrtaka. Mér finnst umdeilanlegt að í því tilviki sé meðaltal hópana sex (5,44) besta spáin fyrir þýðið og tel að miðgildið væri meira sannfærandi. Ástæðan er sú að lægsta meðaltalið (4,17) hefur á sér ákveðinn „útlagablæ“ (og raunar það hæsta einnig). Þannig eru fimm meðaltöl eru hærri „meðaltal meðaltalanna“ en aðeins eitt fyrir neðan það.

Í þessu sambandi má t.d. benda á að í ís-listdansi á ólympíuleikunum þá gefa dómarar upp einkunnir en hæsta og lægsta einkunn eru ekki taldar með þegar meðaltalseinkunn er reiknuð út.

Og ef út í það er farið, hvers vegna er ekki yfirleitt skynsamlegra að miða við miðgildið fremur en meðaltalið í ályktunartölfræði?

Þetta er rétt athugað að lægsta meðaltalið á glæru 4 í *Grunnatriði ályktunartöl-fræði* er áberandi lægra en hin meðaltölin. Ef við tökum meðaltölin sex þá getum við reiknað staðalfrávik þeirra (þ.e. spágildi staðalvillu meðaltalanna) en það er 0,672. Með þessar upplýsingar að vopni sjáum við að lægsta meðaltalið er tæplega einni staðalvillu (staðalfráviki) fyrir neðan meðaltal meðaltalanna. Almennt séð viljum við sjá meira frávik frá meðaltalinu (meðaltali meðaltalanna) áður en við tölum um frávilling (útlaga).

Það má einnig nálgast þetta úr annarri átt. Frávillingar (útlagar) eru sjaldan nákvæmlega skilgreindir. Þá má ýmist skilgreina sem fráviksgildi, þ.e. gildi sem er langt frá öðrum mæligildum. Samkvæmt ofangreindu er lægsta meðaltalið ekki frávillingur í þessari merkingu. Önnur og þróaðri skilgreining byggist á því að frávillingur komi úr öðru þýði heldur en hin mæligildin. Samkvæmt því er lægsta meðaltalið ekki heldur frávillingur (útlagi) því við vitum mæta vel úr hvaða þýði meðaltölin koma og að þau koma öll sex úr sama þýðinu.

Varðandi miðgildið þá kemur í ljós að meðaltalið er besta mælitala á miðleitni (miðsækni) ef þýðið er normaldreift. Í því felst að engin mælitala hefur minni úrtakadreifingu en meðaltalið og t.d. hefur miðtalan meiri úrtakadreifingu. Það er ein meginástæðan fyrir því að ályktunartölfræðin byggir á meðaltölum.

Engin einhlít lausn er á því hvernig eigi að bregðast við því ef þýði er langt frá því að vera normaldreift, aðferðin sem skautadómarar nota er t.d. aðferð sem er velþekkt innan tölfræðinnar. Markgildissetningin bjargar þó mjög miklu; jafnvel þótt þýðið sé ekki normaldreift verður úrtakadreifing meðaltalanna normaldreifð ef úrtökin eru nægilega stór. Um það má sannfærast með því að leika sér með javaforrit sem sýnir áhrif þýðisdreifingar og úrtaksstærðar á úrtakadreifingu ýmissa úrtakstalna.

Hægt er að átta sig á eiginleikum meðaltala og miðgilda með því að leika sér með javaforrit sem gefur færi á að leika sér með áhrif lögunar dreifingarinnar á meðaltöl og miðgildi. Ýmsa aðra tölfræðilega herma er hægt að nálgast á sama stað.

1999-02-22 GBA
1999-02-23 GBA

Marktekt og núlltilgátur

Í síðast hlutaprófi var spurt í einni fjölvalsspurningu hvort „tilgátan mín væri marktæk.“ Hvað er átt við þegar talað er um „tilgátuna mín,“ er það aðaltilgátan?

Ég reyni að forðast að tala um marktækar tilgátur eða marktækan mun meðaltala. Ég kannast ekki við þetta orðalag úr prófinu né finn ég það þegar ég leita.

Tilgátur eru réttar eða rangar en tölfræðipróf eru marktæk eða ómarktæk. Marktækt próf þýðir að niðurstaðan hafi verið ólíkleg ef núlltilgátan er rétt. Ómarktækt próf felur í sér að niðurstaðan sé líkleg ef núlltilgátan er rétt.

Ég kannast ekki við orðalagið „tilgátan mín“ en venjulega er vísað til aðaltilgátu nema sérstaklega sé tiltekið (eða það sé ljóst af samhengi) að talað sé um núlltilgátuna. Aðaltilgátan er sú tilgáta sem þú setur fram og er í þeim skilningi tilgátan þín. Þegar talað er um að rannsakandi hafi sett fram tilgátu eða prófað tilgátu er undantekningarlítið átt við aðaltilgátuna.

1999-04-23 GBA

Gefum okkur að núlltilgátan hafi ekki verið studd í ofangreindri fjölvalsspurningu og að þar með hafi tilgátan mín verið studd, átti ég þá að svara því að tilgátan mín (aðaltilgátan) hafi staðist? Má ég það?

Ef þú *hafnar* núlltilgáti miðað við eitthvert α -stig, þá tekurðu upp aðaltilgátuna. Þú lítur sem sé svo á að aðaltilgátan sé rétt, þ.e. að hún lýsi þýðinu.

Ef aðaltilgátan er tvíhliða, t.d. að líkamshæð karla sé önnur en líkamshæð kvenna, þá ertu formlega búin að draga þá ályktun að meðalhæð kynjanna sé ekki sú sama. Túlkunin á slíkri tvíhliða niðurstöðu fer hins vegar einnig eftir niðurstöðunum í úrtakinu. Ef meðalhæð karla er meiri en kvenna í úrtakinu, þá mættir þú álykta að karlar séu að jafnaði hærri en konur (þýðismeðaltal karla sé hærra en þýðismeðaltal kvenna). Taktu eftir því að heildarályktunin er einhliða, þó svo að aðaltilgátan sé tvíhliða.

1999-04-23 GBA

Gengur núlltilgátan ekki alltaf út á það að ekki hafi átt sér stað nein veruleg breyting, að breytingin falli innan marktektarstigsins (sem ég setti í upphafi)

Þetta er rangt orðað! Tilgátan getur aldrei falið í sér marktekt af neinu taki. Því getur núlltilgátan ekki falið í sér að „breytingin falli innan marktektarstigsins.“ Mundu að aðeins tölfræðipróf eru marktæk eða ómarktæk en tilgátur eru réttar eða rangar. Í þýði er munur milli meðaltala (stundum er enginn munur) og tengsl milli breyta (stundum engin tengsl).

Hefðbundnar núlltilgátur fela í sér „núll“ þ.e. að enginn munur eða tengsl séu í þýði. Það er hugsanlegt að setja fram núlltilgátu sem felur í sér engin „veruleg tengsl“ en það væru ekki hefðbundnar núlltilgátur eins og þær sem fjallað er um í Aðferðafræði II eða sambærilegum námskeiðum.

1999-04-23 GBA

Tilgátuprófun og marktekt

Þú sagðir í tíma um daginn að „Ef ég tel líklegt að núlltilgátan sé rétt þá þarf minna p til að hafna núlltilgátunni.“ Við skiljum þetta ekki!!

Okkur finnst að þetta eigi að vera öfugt, þ.e. ef núlltilgátan er röng þá þurfi minna p. Myndirðu vera svo elskulegur að svara þessu fyrir okkur?

Minna p merkir þarna að SPSS prenti út lægra marktektargildi, t.d. 0,001 í stað 0,05. Þ.e. ég þarf það sem Agresti kallar „*more evidence*“ til þess að hafna núlltilgátunni.

Ef ég trúi ekki á miðla (T_0: Stefán er ekki miðill. T_1: Stefán hefur miðilshæfileika) þá þarftu að leggja hart að þér (gera hluti sem eru mjög ólíklegir undir núlltilgátunni, þ.e. eitthvað sem er ólíklegt ef þú ert ekki miðill; færa hluti úr stað, lækna blinda, o.s.frv.) til að sannfæra mig um að þú sért í rauninni miðill (að T_0 sé röng).

Vor 98 GBA

Ég að velta fyrir mér hvað marktekt er ég hef bara fundið hvað hún er ekki?

Um marktekt er fjallað í köflum 6.2-6.4 í Agresti. Það er erfitt að finna texta sem segir aðalatriðið í stuttu máli heldur þarftu að liggja aðeins yfir honum. Skoðaðu samt það sem Agresti & Finlay segja um þetta á bls. 173 undir millifyrirsögninni „Interpreting the *p*-value.“ Í hrárri þýðingu á íslensku segir þar:

„Þegar *p*-gildið er lágt eru gögnin í andstöðu við núlltilgátuna; niðurstaðan („the observed data;“ gögnin sem við höfum) væri ólíkleg ef núlltilgátan væri rétt.

Algeng mistök eru að rangtúlka *p*-gildið sem líkur þess að núlltilgátan sé rétt. ...“

Svo mörg eru þau orð. Ef tölfræðipróf er marktækt, þá erum við með niðurstöðu sem væri mjög ólíkleg ef núlltilgátan væri rétt. Slíkt veikir traust okkar á núlltilgátunni og leiðir til þess að við höfnum henni.

Vor 98 GBA

Hvenær er núlltilgátu hafnað?

Hvernig veit ég að ég get hafnað núlltilgátu? Sbr. dæmi 55 þar sem $t(19)= 2{,}093$ samkvæmt töflu en útreikningur sýnir að $t(19)= 2{,}186$ og sbr. dæmi 53 þar sem z-taflan segir að $z= 1{,}96$ en útreikningur er $z= -2{,}083$.

Af hverju er hægt að hafna núlltilgátu í báðum þessum tilvikum?

Við tilgátuprófun notarðu ályktunarreglu sem segir að ef niðurstaðan í úrtakinu er ólíkleg ef núlltilgátan væri rétt þá sé þér heimilt að ganga út frá því að núlltilgátan sé röng. Þessi ályktunarregla getur auðvitað brugðist, þ.e. núlltilgátan getur verið rétt þótt reglan gefi til kynna að henni megi hafna en einnig röng þótt reglan segi að þú getir *ekki* hafnað henni.

Næsta skref er að ákvarða hvort niðurstaðan í úrtakinu sé líkleg eða ólíkleg miðað við að núlltilgátan sé rétt. Til að komast að því reiknar þú ýmist *t*- eða *z*-próf. Ef miðað er við tvíhliðapróf (einhliða próf er lítillega öðru vísi) þá skiptir tölugildi niðurstöðunnar máli; þú skoðað því niðurstöðu prófsins óháð því hvort það er plús eða mínus fyrir framan niðurstöðutöluna. Því hærra sem tölugildið er því minni líkur eru á þetta miklu fráviki (eða meira fráviki) frá því gildi sem núlltilgátan gefur upp.

Gerum t.d. ráð fyrir eftirfarandi núlltilgátu; T_0: $\mu_1 = \mu_2$. Samkvæmt henni eru þýðismeðaltölin tvö þau sömu. Ef úrtaksmeðaltölin eru einnig þau sömu verður niðurstaða bæði *t*- og *z*-prófs nákvæmlega 0,0. Því meira sem mismunur úrtaksmeðaltalanna víkur frá þessu því hærra verður tölugildið fyrir niðurstöður marktektarprófsins. Mismunur úrtaksmeðaltalanna gæti t.d. verið 4,6 eða –12,4; seinni mismunurinn er meira frávik frá núlltilgátunni og gefur því hærra tölugildi með t-prófi (nú eða z-prófi eftir því hvort prófið á við). Hér skiptir ekki máli að seinna frávikið er neikvætt svo fremi að tilgátan sé tvíhliða.

Það sem ég þarf að vita er hversu líklegt viðkomandi frávik frá núlltilgátunni er miðað við að núlltilgátan sé rétt. Þetta fæ ég að vita með því að fletta upp í viðkomandi töflu, þ.e. *t*- eða *z*-töflu eftir því hvort prófið er reiknað. Talan í töflunni, vendigildið, gefur til kynna hve hátt tölugildi prófsins þarf að vera til að líkindin á þetta miklu eða meira fráviki verði aðeins 5% (eða 1% eftir því hvort alfagildið er valið). Ef tölugildið er lægra en vendigildið eru meira en 5% (1%) líkur á þetta miklu fráviki; ef tölugildið er hærra en vendigildið, þá eru minna en 5% (1%)

líkur á þetta miklu (eða meira) fráviki frá núlltilgátunni. Athugaðu að vendigildið miðast alltaf við það að núlltilgátan sé rétt.

Ef þínar tölur eru réttar, þá fékkst þú niðurstöðuna $t(19)$=2,186 þegar vendigildið var $t(19)_{0,025}$= 2,093. Niðurstaðan er því hærri en vendigildið og því eru minni en 5% líkur á þetta miklu eða meira fráviki frá núlltilgátunni miðað við tvíhliða próf og miðað við að núlltilgátan sé rétt. Þetta væri því ólíkleg niðurstaða ef núlltilgátan væri rétt og því er þér heimilt að hafna henni.

Á sama hátt fékkst þú z= –2,083 þegar vendigildið var 1,96. Tölugildi niðurstöðunnar (2,083) er hærri en vendigildið og því eru minna en 5% líkindi á niðurstöðunni í úrtakinu (eða meira fráviki frá T_0) og því er þér heimilt að hafna núlltilgátunni og gera ráð fyrir að hún sé röng og aðaltilgátan rétt.

2001-04-10e GBA

Hvað er vendigildi

Er það rétt hjá mér að vendigildi sé eins konar mörk marktektar. Ég á við hvort það sé ekki eins konar punktur þar sem niðurstaðan nær α. Ef niðurstaðan nær vendigildinu, þá má hafna núlltilgátunni en annars megi taka hana upp.

Í aðalatriðum er þetta rétt hugsað. Ef niðurstaða marktektarprófs nær vendigildinu, má hafna núlltilgátunni en annars ekki. Vendigildið skilgreinir því mörkin milli þess að mega og mega ekki hafna núlltilgátunni. Gættu þó að eftirfarandi tveimur atriðum.

Það er ekki mögulegt að *taka upp* núlltilgátuna. Þetta er útskýrt í **Ályktunum í einum hópi** á glærunni **Takmarkanir tilgátuprófunar**.

Það getur verið villandi að tala um að niðurstaðan nái α. Niðurstaða marktektarprófs er t.d. z, t eða χ^2. Ef niðurstaðan nær vendigildinu eru líkurnar á þetta miklu eða meira fráviki frá núlltilgátunni nákvæmlega α. Við getum því sagt að vendigildið sé sú niðurstaða marktektarprófs sem samsvari því að p-gildi niðurstöðunnar sé nákvæmlega α.

Ef vendigildi fyrir einhliða t-próf er t.d. 1,729 (miðað við 95% öryggi og 19 frígráður), þá er mér stætt á að hafna núlltilgátu ef útkoma t-prófsins er hærri en þetta vendigildi (sbr. glæruna **Eykst þol við líkamsrækt** í **Ályktanir í einum hópi**)? Ástæðan væri sú að það væru minna en 5% líkur á að fá þessa útkomu ef núlltilgátan væri rétt?

Já. Það er þó örlítil ónákvæmni í þessu sem vert er að leiðrétta. Líkurnar á þessari *tilteknu* niðurstöðu geta verið mjög litlar. Það sem niðurstaðan segir er að líkurnar á niðurstöðunni eða *meira* fráviki frá núlltilgátunni ef hún er rétt er minna en 5%, þ.e. það α sem var valið í þessu tilviki.

Í tvíhliða prófi þar sem vendigildið væri með sömu forsendum og að ofan 2,09 þá get ég ekki hafnað núlltilgátunni ef niðurstaða t-prófsins er 1,91 þar sem það eru meira en 5% líkur á að fá þessa útkomu ef núlltilgátan er rétt (sbr glæruna **Útreikningur t-prófs í einum hópi**)? Er þetta rétt skilið hjá mér?

Já, en rétt er að orða það eilítið öðru vísi sbr. athugasemd hér fyrir ofan. Það eru meira en 5% líkindi á að fá *þetta mikið eða meira* frávik frá núlltilgátunni í aðra hvora áttina ef núlltilgátan er rétt.

Það væri kannski gott að fá líka svart á hvítu nákvæmlega hvað vendigildi er þ.e. skilgreiningu svo ég nái þessu nú í eitt skipti fyrir öll.

Ágæta skilgreiningu má fá í Hyperstat Online og aðra hjá Statistics Glossary.

2002-04-29a GBA

p-gildi, *α* og vendigildi

Ég er að velta fyrir mér spurningu 64 í Stoðkverinu. Svarið við dæminu er $t(69)= 4{,}8$, $p < 0{,}001$. Ég hefði haldið að það ætti að vera $p < 0{,}05$ þar sem miðað er við $\alpha= 0{,}05$. Af hverju 0,001 og hvernig er þetta fengið?

Það eru í aðalatriðum tvær aðferðir við að ákvarða marktekt. Í báðum tilvikum vel ég α sem ég vil miða við þegar ég ákvarða marktekt.

Í öðru tilvikinu fletti ég upp vendigildi, sbr. fyrirspurnina **Hvað er vendigildi** í Spurðu og svöruðu. Ef tölugildi niðurstöðu prófsins er hærri en vendigildið, þá hafna ég núlltilgátunni og segi prófið vera marktækt miðað við það α sem ég gaf upp (t.d. 0,05 eða 0,01).

Hin leiðin er að reikna út nákvæmt *p*-gildi. Þá reikna ég út nákvæmt *p*-gildi og ef það er lægra en α hafna ég núlltilgátunni og lýsi prófið marktækt. En þar sem ég er með nákvæmt *p*-gildi í höndunum, get ég gefið það upp en ekki eingöngu það α sem ég miðaði við.

Það eru að skiptar skoðanir um hvora leið skuli fara. Seinni leiðin er þó nánast allsráðandi enda miða öll tölfræðiforrit við hana. Oftast er þó ekki gefið upp nákvæmt *p*-gildi heldur miðað við nokkur stöðluð gildi svo sem 0,05, 0,01, 0,005 og 0,001.

Nákvæmt *p*-gildi get ég ekki reiknað nema með forriti, en þessi stöðluðu gildi get ég oftast lesið úr töflum með því að finna vendigildin fyrir hvert og eitt þeirra. Niðurstaðan er því í reynd borin saman við nokkur vendigildi.

2003-05-01a GBA
2003-12-29a GBA

Útreikningur á afköstum og fastheldnimistökum

Mig langaði til að vita hvort við þyrftum að geta reiknað afköstin í tilgátuprófunum eða hvort okkur væru gefnar upp tölur fyrir Beta mistökin þannig að við reiknum þá 1–*P*(beta mistök) og fengjum þá afköstin.

Eftir að vera búin að glugga mikið í A 6.7 þá finnst mér ég geta áttað mig svolítið á hvernig á að reikna afköstin en þó óttast ég að gera villur í prófinu.

Þú munt ekki þurfa að reikna afköstin í prófinu. Hins vegar geturðu þurft að skilja hvað í þeim felst og gefa til kynna skilning á tengslunum milli fastheldnimistaka (β) og afkasta (*power*), þ.e. $1-\beta$.

Vor 1999 GBA

Hvernig ráðast líkur fastheldnimistaka af staðalvillu?

Líttu fyrir mig á glæruna Rökfræði tilgátuprófunar í Ályktunum í einum hópi. Hvernig ráðast líkur fastheldnimistaka af staðalvillunni? Stendur β fyrir staðalvillu í töflunni á þessari glæru ?

Setjum sem svo að viðmiðstalan er 50,2 þ.e. þýðismeðaltalið undir núlltilgátunni. Þá ættu úrtaksmeðaltölin, ef núlltilgátan er rétt, að falla einhvers staðar í kringum 50,2. Ef staðalvillan er lítil falla úrtaksmeðaltölin nálægt viðmiðsgildinu (undir núlltilgátunni) en gætu fallið langt frá ef staðalvillan er stór.

Setjum sem svo að úrtaksmeðaltalið sem þú færð sé 54,0. Ef staðalvillan er lítil þá væri ólíklegra að fá þetta hátt meðaltal en ef staðalvillan er stór. Þú ert því líklegri til að hafna núlltilgátunni með litla staðalvilla heldur en með stóra. Ef þýðistalan *er* 54,0, þá er núlltilgátan auðvitað röng og þú vilt hafna henni. Auknar líkur á að hafna rangri núlltilgátu minnka líkur á fastheldnimistökum (Fastheldnimistök er það að hafna ekki núlltilgátu þegar hún er röng).

Lítil staðalvilla þýðir því lágt p-gildi og mikil marktekt. Há staðalvilla þýðir hátt p-gildi og lítil marktekt. Ef núlltilgátan er röng vildu háa marktekt; lítil staðalvilla eykur líkur þess.

Ef núlltilgátan er *rétt* skiptir staðalvillan auðvitað ekki máli fyrir marktekt. Líkurnar á marktekt verða α (t.d. 0,05) hver sem staðalvillan er. Það er eingöngu þegar núlltilgátan er *röng* sem staðalvillan hefur áhrif á niðurstöður.

Beta (β) eru líkur á því að hafna ekki rangri núlltilgátu. Þú minnkar β með því að minnka staðalvilluna. Samt er β ekki það sama og staðalvillan.

Stúderaðu vandlega bls. 193–195 í Agresti.

2000-03-17 GBA

Einhliða og tvíhliða próf

Hvernig sér maður muninn á því hvenær maður notar tvíhliðapróf eða einhliðapróf?

Það ræðst af spurningunni sem er til umfjöllunar. Lestu vel verkefnin í Stoðkverinu og æfðu þig í því að ákvarða hvort tilgátan sé með stefnu eða stefnulaus.

Í verkefni 84 í Stoðkveri er spurningin er orðuð þannig að eðlilegast er að framkvæma einhliða próf. Væri það rangt á prófi að leysa slíkt dæmi sem tvíhliða próf, þ.e. mundi það reiknast til frádráttar?

Já, það kæmi til frádráttar. Einkunn myndi lækka lítillega vegna þessa.

2003-01-11g GBA

Tilgátur í verkefni 79 í Stoðkveri

Ég er með fyrirspurn vegna dæma í stoðkveri. Ég er ekki viss í dæmi 79 hverjar tilgáturnar eru; hver er núlltilgátan og aðaltilgátan?

Hvernig er best að sjá hverjar tilgáturnar eru og hvernig þær eru settar fram (hver er núlltilgáta og aðaltilgáta) í dæmum sem þessum? Það er jú forsendan fyrir því hvernig maður reiknar úr þeim og ályktar.

Tilgáturnar ráðast af rannsóknarspurningunni. Í þessu tilviki er verið að athuga hvort niðurstöður prófsins séu aðrar en erlenda staðalbindingin gefur til kynna. Aðaltilgátan er því að meðaltalsniðurstaðan hérlendis sé önnur en í erlendu staðalbindingunni, þ.e. ekki 10,0. Núlltilgátan er andhverfa þessa, þ.e. að þýðismeðaltalið sé nákvæmlega 10,0.

Í dæmi 79 kemur þetta skýrt fram í lið b, þar sem segir að gera ætti ráð fyrir að þýðismeðaltalið væri 10,0. Sú tala er því viðmiðsgildið. Tilgangur rannsóknarinnar er að kanna hvort prófið mæli öðru vísi hérlendis heldur en erlendis og því er aðaltilgátan að þýðismeðaltalið sé ekki 10,0.

2002-03-15a GBA
2003-01-11b GBA

Hvað á að gera í verkefnum 80 og 81 í Stoðkveri?

Í dæmi 80 og 81 þá er ég ekki viss hvað það er sem ég á að gera; ég næ ekki að lesa úr spurningunni. Eru einhver ákveðin atriði sem hægt er að leita að í spurningunni?

Í þessum verkefnum er ekki lýst nákvæmlega hvaða aðferðir eigi að nota. Þetta er í samræmi við þann veruleika sem mætir okkur þegar námskeiðinu lýkur. Það eina sem er tilgreint eru þær spurningar sem við viljum fá svarað.

Liður a í verkefni 80 biður t.d. um upplýsingar um meðalframmistöðu íslenskra forskólabarna. Þessu þarf að svara með aðferðum ályktunartölfræðinnar, þ.e. með bilspá, punktspá eða tilgátuprófun.

Hér er og í verkefni 81 er því gerð sú krafa að þú getir valið viðeigandi úrvinnsluaðferð í samræmi við þá spurningu sem er til athugunar.

Ég vil benda á svörin fyrir þá sem eru alveg strand. Þar eru gefnar niðurstöður og greint frá þeim aðferðum sem eðlilegast er að nota til að svara viðkomandi spurningum.

2002-03-15b GBA