

Bætt tölfræðinám við Háskóla Íslands

Lokaskýrsla
(Skýrsla vegna styrks fyrir 2021)

Útdráttur

Ráðgert var að semja æfingar þar sem nemendur myndu æfa sig í notkun og túlkun tölfræðilegra hugtaka. Um eitt þúsund slíkar æfingar voru búnar til og settar upp í opna vefkerfinu tutor-web. Æfingarnar innihéldu slembitölur og prófuðu skilning á túlkun niðurstöðu aðhvarfsgreininga. Aðferðafræði æfinganna byggði á niðurstöðum rannsókna, sem birtar voru á ráðstefnunni INTED 2021. Fyrstu niðurstöður greininga á notkun þessarar aðferðafræði benda eindregið til þess að aðferðin leiði til yfurfæranlegrar kunnáttu, sem er talsvert umfram það sem hefur verið sýnt fram á áður.

Æfingasöfn voru hönnuð, mest til aðstoðar nemendum sem vilja eða þurfa að skilja hugtök í einfaldri línulegri aðhvarfsgreiningu.

Vegna COVID-19 fjarkennslu og fjarprófa voru búin til yfir 2 þúsund tutor-web tölfræðidæmi fyrir námskeiðið Líkindareikningur og tölfræði á vormisseri 2020 (LoT2020), til viðbótar þeim þúsundum dæma sem voru í kerfinu fyrir.

Sumarið 2020 fóru síðan fram ítarlegar greiningar á lærdómi nemenda sem notuðu tutor-web til æfinga í tölfræði, mest LoT2020, en einnig í fjölmörgum öðrum námskeiðum. Niðurstöðum er lýst í nokkrum greinum á INTED 2021 ráðstefnunni (mars 2021, sjá viðhengi 1-3).

Í ljós kom að nemendur læra mikið af því að nota kerfið en nokkuð er um utanaðbókarlærdóm einnig. Lagðar hafa verið fram nokkrar leiðir til að koma algerlega í veg fyrir þann utanaðbókarlærdóm sem barna mældist, í þessum greinum og fyrri greinum.

Almenn sátt er meðal nemenda um notkun tutor-web og löngu búið að sýna fram á að nemendur læra ekki minna af tutor-web æfingum en hefðbundnum heimadæmum. Þrátt fyrir þetta er nokkuð um að nemendur nái hárrí einkunn í tutor-web (og öðrum heimaverkefnum) en sýni síðar skort á skilningi þegar kemur að prófi.

Æfingarnar eru aðgengilegar öllum kennurum við HÍ. Kerfið er fyrst og fremst til æfinga og nemendur fá með því símat ásamt endurgjöf eftir hverja æfingu. Yfirferð er sjálfvirk og kennari getur notað einkunn úr kerfinu sem hluta af misseriseinkunn.

Árin 2019 og 2020 voru gerðar tilraunir með að semja spurningar sem innihalda slembitölur og jafnvel gagnasöfn sem valin eru slembið úr gagnasöfnum Hagstofunnar. Með þessu móti er hægt að biðja nemendur um að túlka niðurstöður tölfræðiprófs eða aðhvarfsgreiningar án þess að þurfa að framkvæma greininguna og án þess að hafa séð sömu tölur áður.

Nýja spurningasafnið

Í þessu verkefni var lagður grunnur að því að semja almennar spurningar um tölfræðileg efni. Áherslan í fyrstu atrennu var á túlkun niðurstaðna úr línulegri aðhvarfsgreiningu. Lesandinn getur prófa æfingarnar með því að fara á hlekkinn <https://tutor-web.net/math/math612.0/lecture530>, velja þar “Take a drill on this lecture”, skrá sig inn sem “gestur” með lykilorð “gestur” og velja “Subscribe to this lecture” ef með þarf. Þá hleðst spurningar safn í vafra notandans og hægt að biðja um nýja spurningu.

Hver spurning hér samanstendur af “haus” og valmöguleikum.

Dæmigerður haus er hér til hægri, en hann sýnir úttak úr línulegri aðhvarfsgreiningu með forritinu R.

A student has collected data and run it through the R regression package (lm and summary) to generate the following output:

```
Call:
lm(formula = y ~ x)

Residuals:
Min      1Q  Median      3Q      Max
-6.9198 -1.6821 -0.2686  1.3058  6.5449

Coefficients:
(Intercept)  -1.2999      2.7774  -0.468  0.65225
x              1.5970      0.4476  3.568  0.00732 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.066 on 8 degrees of freedom
Multiple R-squared:  0.6141,    Adjusted R-squared:  0.5658
F-statistic: 12.73 on 1 and 8 DF,  p-value: 0.007318
```

Sem dæmi um svarmöguleika má nefna listann hér til hægri. Hér eru notaðar allar tölurnar úr aðhvarfsgreiningartöflunni, en aðeins ein fullyrðing er rétt.

- a. The value of the t-statistic for testing whether the slope is equal to zero is 10
- b. The correlation is 3.5677
- c. The p-value for testing whether the slope is equal to zero is 8
- d. R-squared= 0.7836
- e. The percentage of variability explained by the model is -1.2999
- f. The std.error of the slope is 3.5677
- g. The std.error of the slope is 0.4476
- h. The p-value for testing whether the intercept is equal to zero is 2.7774
- i. The value of the t-statistic for testing whether the slope is equal to zero is 61.4058
- j. The estimate of the slope is 8
- k. The std.error of the slope is -0.468
- l. The value of the t-statistic for testing whether the intercept is equal to zero is 0.7836
- m. The p-value for testing whether the intercept is equal to zero is 10

Annað dæmi um spurningagerð er hér til hægri. Hér eru gögnin sýnd auk niðurstöðunnar.

A regression analysis was undertaken to estimate the coefficients and properties of the model $Y_i \sim n(\alpha + \beta x_i, \sigma^2)$.

The data and output from R are given in the following:

```
i x y
1 0.5 3.92
2 1.0 4.91
3 1.5 4.00
4 2.0 7.03
5 2.5 6.74
6 3.0 6.77
7 3.5 8.05
8 4.0 9.20
9 4.5 9.44
```

```
> summary(lm(y~x))
```

```
Call:
lm(formula = y ~ x)
```

```
Residuals:
```

```
Min      1Q   Median       3Q      Max
-1.24700 -0.08600  0.06667  0.37617  1.06983
```

```
Coefficients:
```

```
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.1075     0.5060   6.142 0.000471 ***
x            1.4263     0.1798   7.932 9.63e-05 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.6964 on 7 degrees of freedom
Multiple R-squared:  0.8999,    Adjusted R-squared:  0.8856
F-statistic: 62.92 on 1 and 7 DF,  p-value: 9.625e-05
```

Aftur eru gefnir margir svarmöguleikar til að gera erfitt að giska á rétta svarið.

- a. The estimate of the slope is 3.1075
- b. The correlation, R , is 0.8999
- c. The intercept is significantly different from zero (at the 5% level) since $0.0000963 < 0.05$
- d. In simple linear regression, the F-statistic for dropping all independent variables from the model is also the t for slope=0, i.e. $F=6.142$
- e. The sample size is 7
- f. The P-value for testing whether the intercept is zero is 0.000471
- g. The P-value for testing whether the slope is zero is 0.000471
- h. The estimated standard error of the intercept is 0.1798

- Notkun slembitalna kemur í veg fyrir að unnt sé að læra utanað réttar tölur.
- Notkun margra svarmöguleika kemur í veg fyrir ágiskun.
- Notuð eru sitt á hvað orð eða tákn til að lýsa hugtökum
- Spurt er um gildi á tilteknum stærðum en líka hvað það þýðir að tiltekið gildi sé hátt eða lágt

Þriðja dæmið um spurningagerð er að leita að kunnáttu um nákvæmlega tiltekna stærð.

Eins og áður inniheldur spurnoinginr slembitölur. Alls voru framleidd 25 gagnasöfn, en aðuvelt væri að hafa þau miklu fleiri.

Anna Helga has collected data and run it through the R regression package (lm and summary) to generate the following output:

```
Call:
lm(formula = y ~ x)

Residuals:
Min      1Q  Median      3Q      Max
-6.2420 -1.7332 -0.3269  2.2761  4.8033

Coefficients:
(Intercept)  5.3912  2.3367  2.307  0.0499 *
x            0.4016  0.3766  1.066  0.3174
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.421 on 8 degrees of freedom
Multiple R-squared:  0.1245,    Adjusted R-squared:  0.01501
F-statistic: 1.137 on 1 and 8 DF,  p-value: 0.3174
```

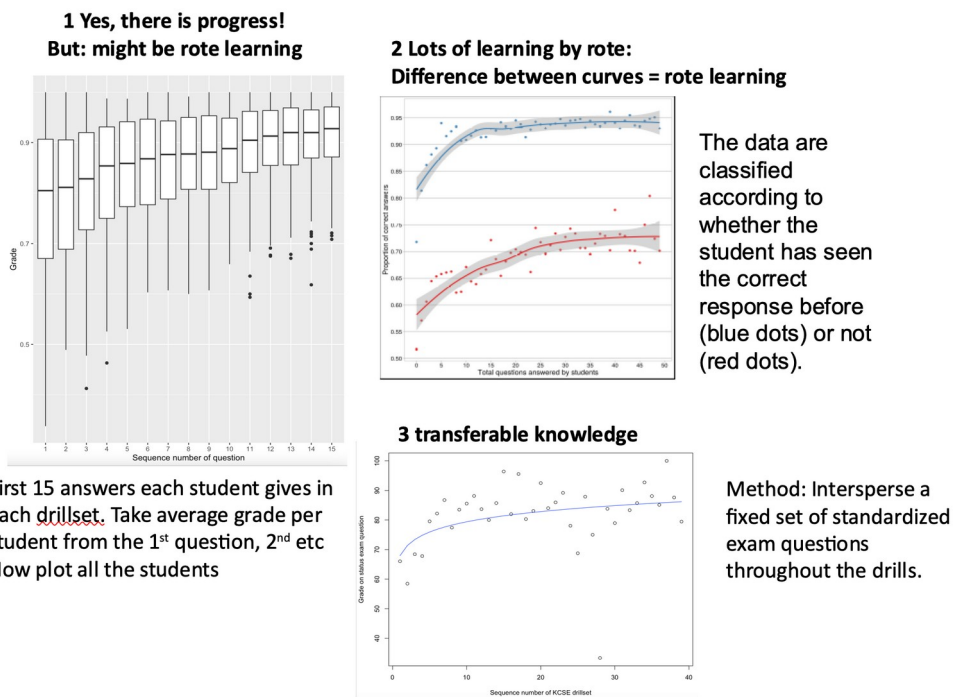
Nú viljum við vita, hvort nemandinn geti lesið gildi metins staðalfráviks úr tölunni.

Aðrar spurningar leita síðan að skilningi á öðrum tölum.

- a. The estimated standard deviation of the response is 3.4206
- b. The estimated standard deviation of the response is 5.3912
- c. The estimated standard deviation of the response is 2.3367
- d. The estimated standard deviation of the response is 2.3072
- e. The estimated standard deviation of the response is 0.0499
- f. The estimated standard deviation of the response is 0.4016
- g. The estimated standard deviation of the response is 0.3766
- h. The estimated standard deviation of the response is 1.0664
- i. The estimated standard deviation of the response is 0.3174
- j. The estimated standard deviation of the response is 0.1245
- k. The estimated standard deviation of the response is 12.445
- l. The estimated standard deviation of the response is 0.3528
- m. The estimated standard deviation of the response is 8

Rannsóknaniðurstöður

Helstu niðurstöður um áhrif mismunandi aðferða við að semja æfingasurningar má sjá í meðfylgjandi mynd.



1. Ef tekið er saman almennt yfirlit til að sýna hvernig nemendum fer fram smátt og smátt innan hvers æfingasafns fæst fyrsti ferillinn á myndinni. Hér er sýnd hvers meðaleinkunn nemanda, innan hvers æfingasafns, teiknað á móti raðtölu spurningarinnar. Augljós galli við þessa aðferð er sá, að þar sem spurningar eru valdar slembið, er ekki hægt að gera greinarmun á einkunn eftir því hvort nemandinn hefur séð spurninguna áður (og gæti þá verið að læra svörin utanað). Þótt tölurnar séu nýjar hafa þessar niðurstöður verið þekktar fyrir tutor-web æfingar síðan 2003.
2. Ef gerður er greinarmunur á einkunnum fyrir fyrsta og annað sinn sem nemandinn sér viðkomandi réttan svarmöguleika sést greinulegur munur og því gæti verið um talsverðan utanaðbókarlærdóm að ræða. Þrátt fyrir það sést einnig að einkunn hækkar smátt og smátt eftir því sem nemandi vinnur lengur innan viðkomandi æfingasafns. Þessar niðurstöður voru birtar 2021 og byggja á gögnum úr tölfræðinámskeiði við HÍ, sem notar tutor-web gögnin, en ekki spurningar af því tagi sem hér voru þróaðar.
3. Síðasta línuritið sýnir nýjar niðurstöður sem byggja á greiningu gagna frá Kenýa 2021. Æfingarnar voru gerðar á sama hátt og þær sem voru gerðar með þeim styrk sem hér er lýst, en ekki eru komin gögn úr íslensku námskeiðunum. Til að taka tillit til þess að nemendur eru misjafnir, æfingarnar eru misjafnar og til að meta almenna aukningu í kunnáttu var eftirfarandi gert: Staðlað safn 18 prófspurninga var sett í heili lagi í öll tutor-web æfingasöfnin (um 50) til stúdentsprófs í stærðfræði í Kenýa. Nemendur fá öðru hvoru slíka spurningu, með tilviljanakenndum hætti. Nemendum er sett það markmið að þau eiga að halda áfram æfingum þar til þau ná 9.75 í einkunn í hverju æfingasafni. Valdir voru úr nemendum sem luku öllum

æfingasöfnum á þennan hátt og höfðu fengið allar spurningarnar. Síðan er meðaleinkunn á æfingasafninu teiknum á móti raðtölu æfingasafnsins. Þessar fyrstu niðurstöður benda til þess að nemendur fá stöðugt hærra einkunn á stöðuprófinu eftir því sem þau æfa sig lengur. Rétt er að benda á, að **spurningarnar á stöðuprófinu eru ekki hluti af æfingasafninu**, Hér eru því vísbendingar um að nemendur öðlist yfirfæranlega þekking við að æfa sig í almennri stærðfræði í vefkerfinu.

Augljóslega þarf að gera sambærilegar greiningar þegar fyrir liggja álíka gögn úr tölfræðinámskeiðum við HÍ.

Heimildir

Á Google scholar má finna 30 greinar um tutor-web og umbunarkerfi þess, sjá td <https://bit.ly/3b66usP>