

Einföldun

Sumarverkefni

Er hægt að mæla framfarir en ekki bara utanaðbókarlærdóm?

1 Inngangur

Aðeins til að byrja, þetta er tilraun hjá mér í að ná að orða hluti aðeins betur og koma gögnunum fram á skiljanlegan hátt. Það vantar líklega enn aðeins upp á það. En það er markmið skýrslunar hér.

Nú, rannsóknarspurninginn sem verið er að vinna með hér er:

Er hægt að mæla framfarir en ekki bara utanaðbókarlærdóm? Bera saman fyrstu svarmöguleika og þá sem hafa komið áður.

2 Gögninn

Til að byrja með, er gott að skoða gögninn. Ég hef fjarlægt sumar breytur, þar sem þær voru ruglingslegar og óþarfi.

```
hashAnswer <- read.csv('Data/hashAnswer4.csv') %>% subset(select = -c(timeStart, fsvfat, fsvfatu, X))
hashAnswer$hsta <- hashAnswer$hsta %>% as.character()
hashAnswer$lectureId <- hashAnswer$lectureId %>% as.factor()
hashAnswer$studentId <- hashAnswer$studentId %>% as.factor()
hashAnswer$nicc <- hashAnswer$nicc %>% as.factor()
hashAnswer %>% glimpse()
```

```
## Rows: 96,018
## Columns: 11
## $ lectureId <fct> 3082, 3082, 3082, 3082, 3082, 3082, 3082, 3082, 3082, 30...
## $ studentId <fct> 1, 1, 18788, 18788, 18788, 18788, 18788, 18788, 18788, 1...
## $ questionId <int> 45591, 44899, 44920, 45021, 45480, 45463, 45101, 44974, ...
## $ correct <int> 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1,...
## $ hash <chr> "d3a978206c877eee8627a29c7a94238d8339d66fc4be460d9e6b6cb..."
## $ fsvfat <int> 0, 1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15,...
## $ hsta <chr> "0", "0", "0", "0", "0", "0", "0", "0", "1", "0", "0", "...
## $ hluta <dbl> 0.0000000, 0.0000000, 0.0000000, 0.1428571, 0.0000000, 0...
## $ timeDif <int> NA, NA, NA, NA, NA, NA, NA, NA, 254, NA, NA, NA, NA, NA,...
## $ nicc <fct> 3, 3, 3, 7, 3, 3, 4, 5, 6, 6, 5, 3, 7, 3, 6, 5, 5, 3, 3,...
## $ gpow <dbl> 8.2506619, 8.2506619, 0.3681775, 0.3681775, 0.3681775, 0...
```

Gagnasettið, eins og það er núna hefur 10 breytur. Breyturnar eru:

- lectureId: Þetta er Id breyta fyrir hvaða fyrirlestur við erum í.
- studentId: Þetta er Id breyta fyrir hvaða nemenda við erum með.
- questionId: Þetta er Id breyta fyrir hvaða spurningu við erum með.
- correct: Þetta er binary breyta sem segir til hvort það var svarað rétt eða rangt.
- hash: Þetta er hash breyta sem segir til hvaða rétta svar við erum með hérna
- fsvfat: “fjöldi spurninga fram að þessu” þetta segir til um hve margar spurningar nemandinn hefur svarað hingað til, innan við þennan fyrirlestur. Talning byrjar á 0.
- hsta: “hef séð þetta áður” er breyta sem segir til um hvort að nemandinn hefur séð rétta svarið áður eða ekki
- timeDif: Þetta er breyta sem segir okkur hve langt það var séðan nemandinn sá þetta svar áður.

- `nicc`: þetta er “number of incorrect choices”, líke þekkt sem `NumQ`. Segir til hve margir rangir svarmöguleikar það eru.
- `gpow`: `gpow` er breyta sem er sett á nemenda þegar hann byrjar á fyrirlestrinum sem segir til hve hratt hann fær erfiðari spurningar.

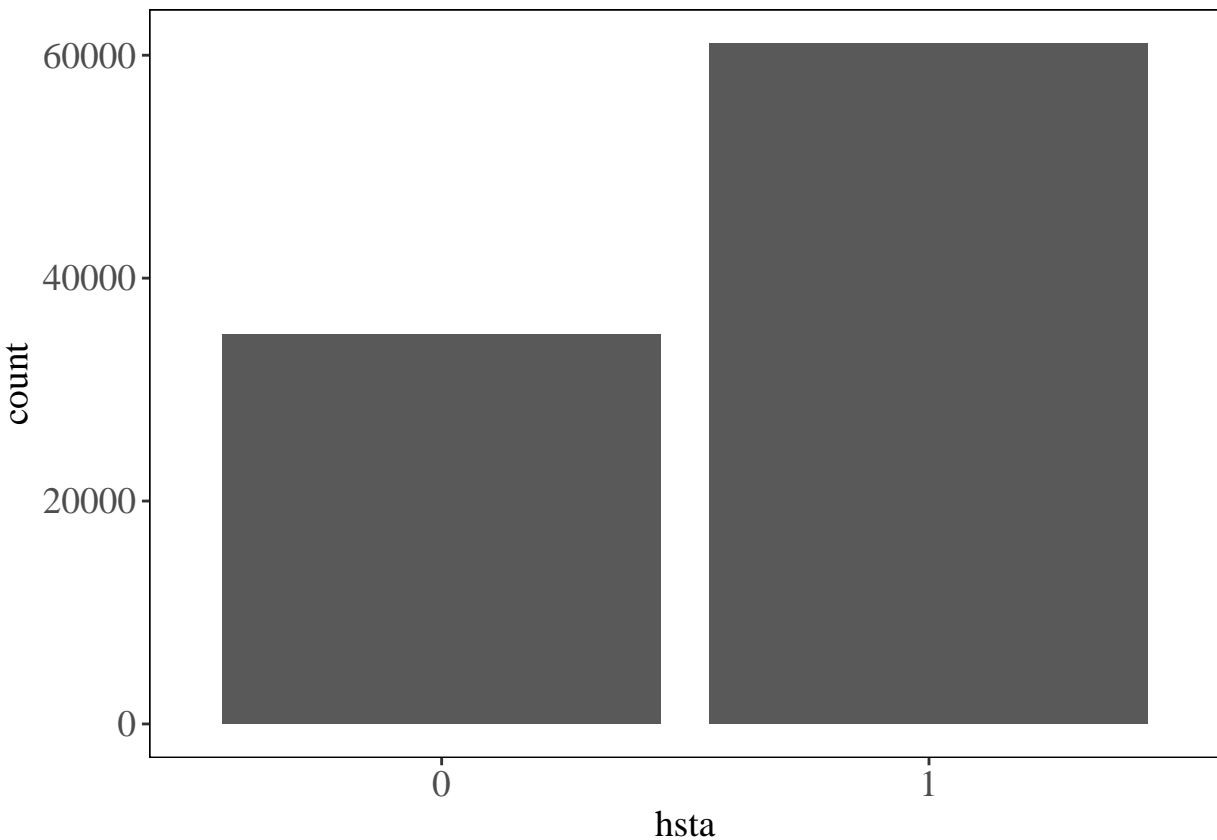
3 Teikningar og töflur

Fyrsta sem er gott að gera, er að skoða aðeins gögninn sjálf.

3.1 skoðað fyrir öll gögninn

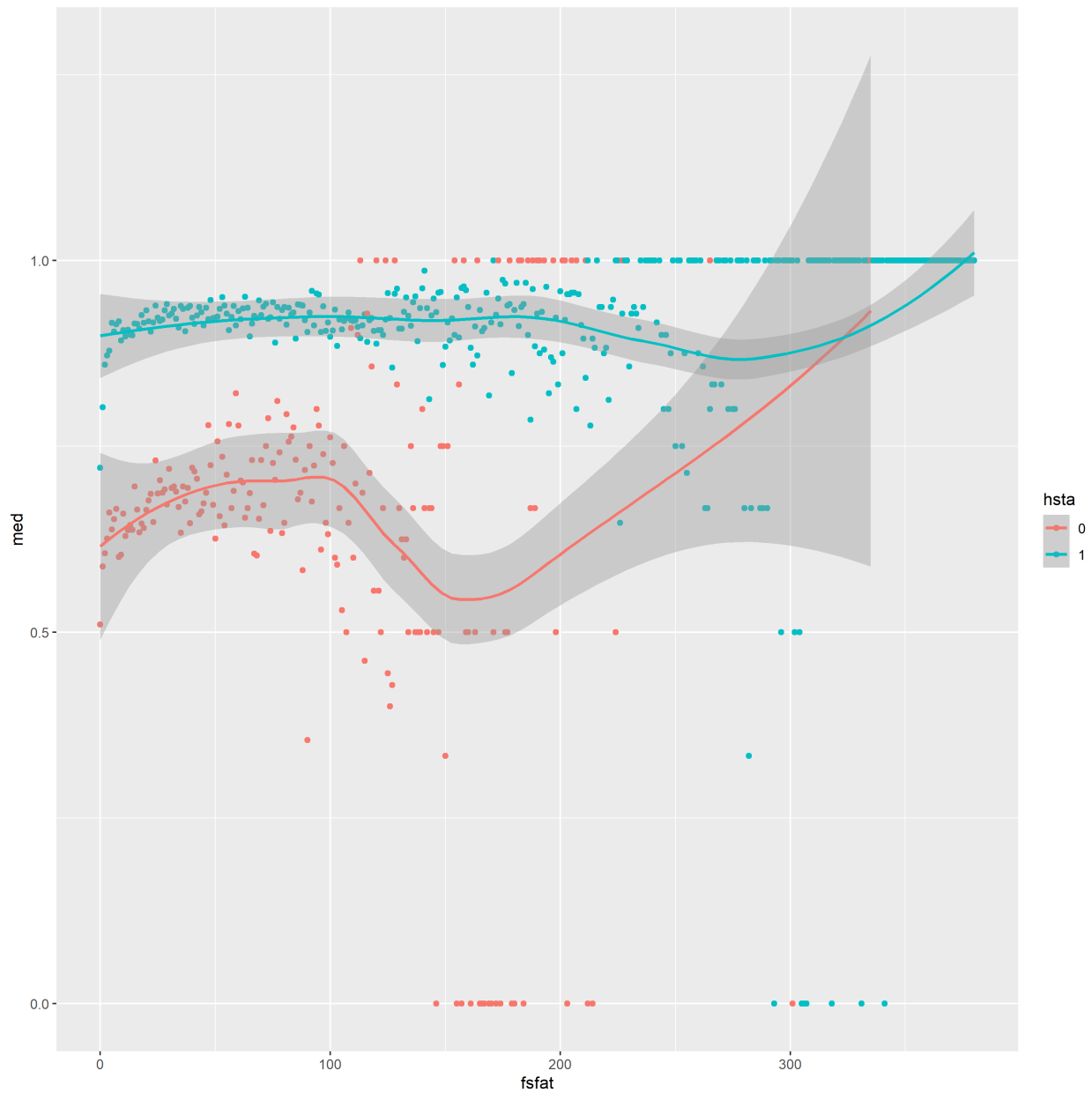
Byrjum aðeins á því að skoða hve mikið af gögnunum eru nemendur að sjá nýtt svar á móti gömlu svari.

```
ggplot(hashAnswer, aes(x = hsta)) +  
  geom_bar()
```



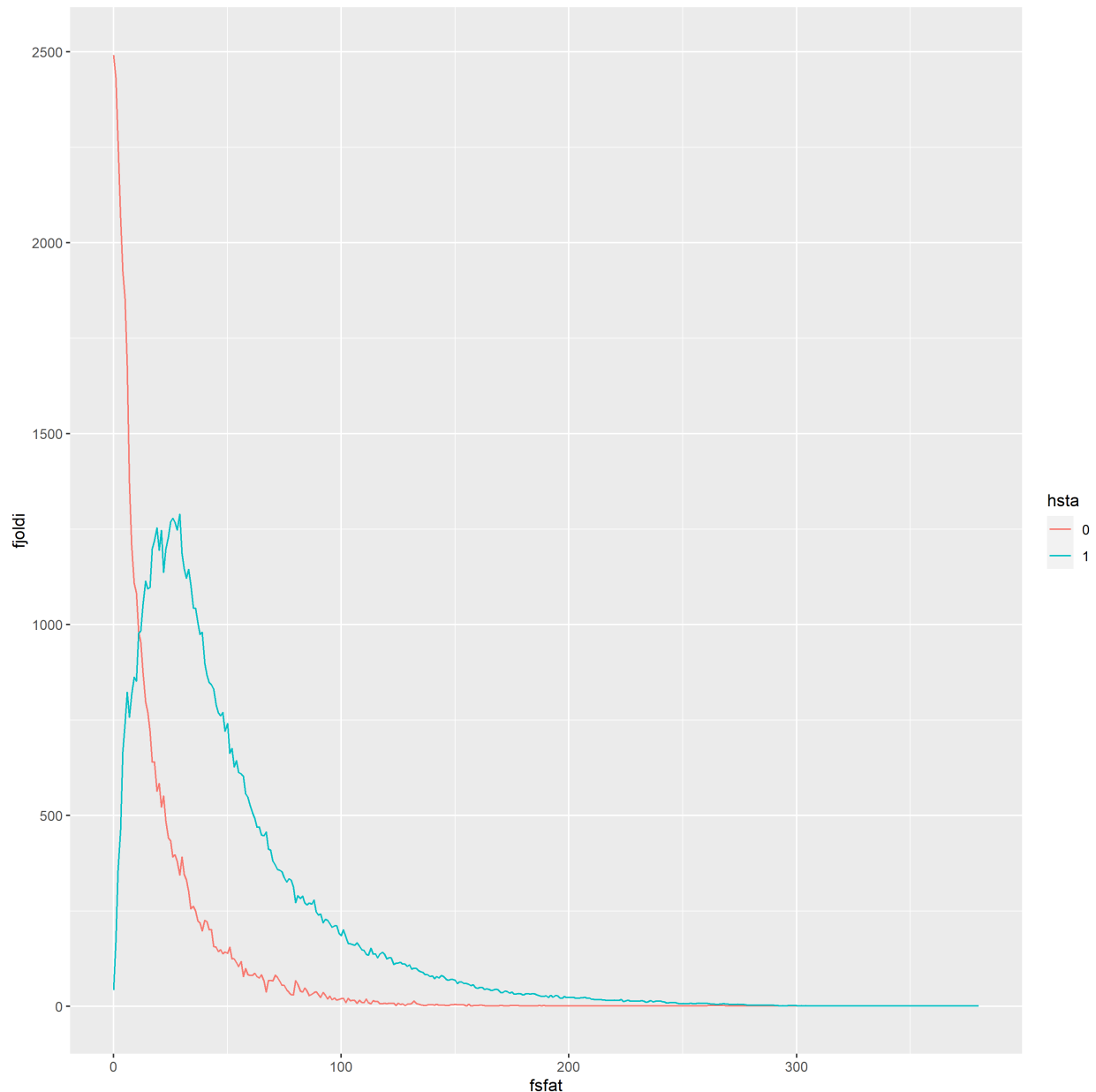
Það lítur út fyrir að vera mikið meira af gögnunum sem eru svör sem hafa sést áður á móti nýjum svörum. Þar sem u.þ.b. 2/3 af gögnunum eru spurningar með rétt svör sem hafa sést áður á meðan 1/3 af gögnunum eru spurningar með rétt svör sem eru að sjást í fyrsta sinn.

Nú markmiðið er að geta skoðað framfarir nemenda, ekki bara utanaðbókarlærdóm, svo það væri gott að setja upp graf sem inniheldur meðaltal réttra svara fyrir hvert fsfa og skipt eftir hsta. Samsagt meðaltal réttra svara hjá þeim sem hafa séð x margar spurningar hingað til, bæði fyrir þann fjölda sem hefur séð rétta svarið áður og þau sem hafa það ekki.



Mynd 1: plot by mean

Það er nú smá erfitt að lesa þessa mynd, það lítur út fyrir að vera smá vaxi hjá þeim sem eru að sjá þetta í fyrsta skiptið, svo fellur það niður og aftur upp. Svo er dreifnin að vaxa því lengra sem er farið. Skoðum aðeins stærð gagnanna fyrir hvern punkt til að sjá hvað er að gerast



Mynd 2: plot by sum

Hérna kemur skýringinn, það lítur út fyrir að vera stórt fall af gögnum sem eru að sjá svar í fyrsta sinn því lengra sem er farið, og líka hjá þeim sem hafa séð svarið áður. Þetta er líklega útaf því að hver nemandi fer mismunandi langt og sumir nemendur geta verið að fara lengur en hinir. Skoðum aðeins hlutföll gagnanna til að skoða þetta betur.

```
a <- hashAnswer %>% summarise("FY50" = sum(fsfat >= 50), "FY100" = sum(fsfat >= 100),  
                              "FY150" = sum(fsfat >= 150), "FY200" = sum(fsfat >= 200),  
                              "FY250" = sum(fsfat >= 250), "FY300" = sum(fsfat >= 300))
```

```

b <- hashAnswer %>% summarise("HY50" = mean(fsfat >= 50), "HY100" = mean(fsfat >= 100),
                              "HY150" = mean(fsfat >= 150), "HY200" = mean(fsfat >= 200),
                              "HY250" = mean(fsfat >= 250), "HY300" = mean(fsfat >= 300))
ab <- cbind(a, b)
FHbylim <- ab %>% pivot_longer(c('FY50', 'HY50', 'FY100', 'HY100', 'FY150', 'HY150', 'FY200', 'HY200',
                              names_to = "typewLim", values_to = "values") %>%
  separate(typewLim, into = c("type", "limit"), sep = 2) %>% pivot_wider(names_from = type, values_from = values)

FHbylim %>% kable(digits = 4)

```

limit	FY	HY
50	28325	0.2950
100	8734	0.0910
150	2970	0.0309
200	1098	0.0114
250	323	0.0034
300	84	0.0009

Lítur út fyrir að um 27.75% gagnanna eru af spurningum sem fara yfir 50 fsfat og bara u.þ.b. 8.29% gagnanna sem fara yfir 100. Hugmynd er þá að skera af gögninn miðað við 50 fsfat eða 100 fsfat. Skoðum þá aðeins hlutfall hsta og meðtalals grafið að ofan miðað við það.

3.2 Skoðað eftir limit

Fyrst er að setja upp gögninn. Að auki fjarlægji ég nemendur sem svöruðu ekki fleiri en 7 spurningum samtals

```

hashLim50 <- hashAnswer %>% group_by(studentId) %>% mutate("count" = n()) %>%
  filter(count > 7 & fsfat < 50)
hashLim100 <- hashAnswer %>% group_by(studentId) %>% mutate("count" = n()) %>%
  filter(count > 7 & fsfat < 100)

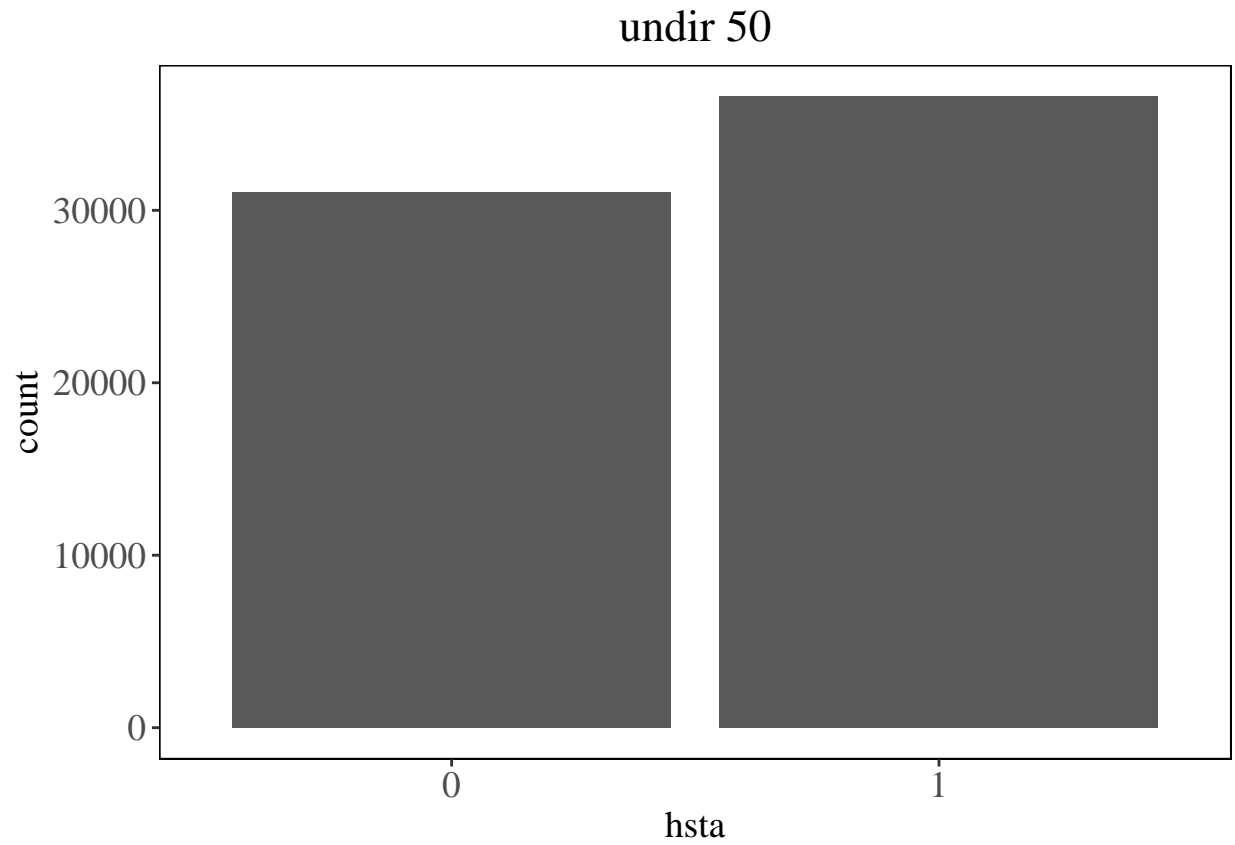
```

Skoðum hlutfallið aftur hjá hsta.

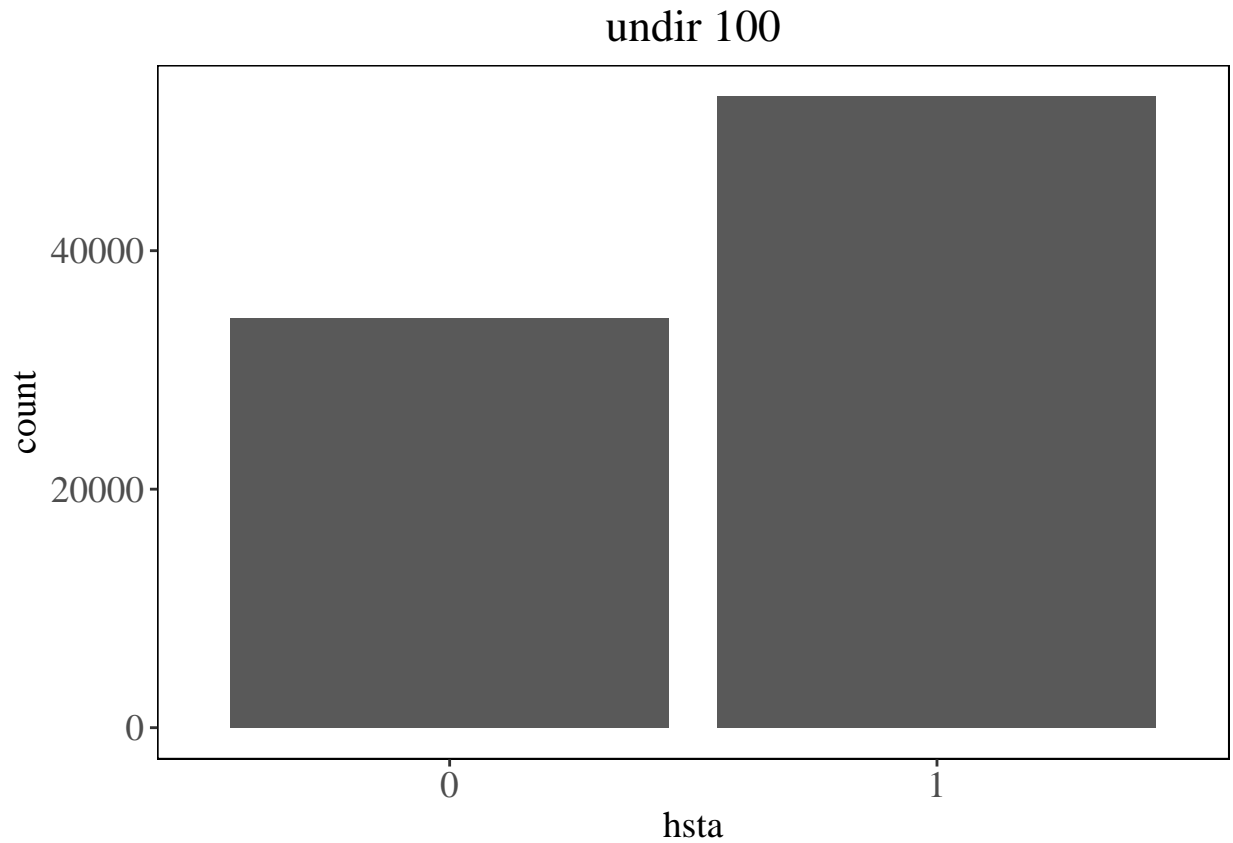
```

ggplot(hashLim50, aes(x = hsta)) +
  geom_bar() +
  ggtitle("undir 50")

```

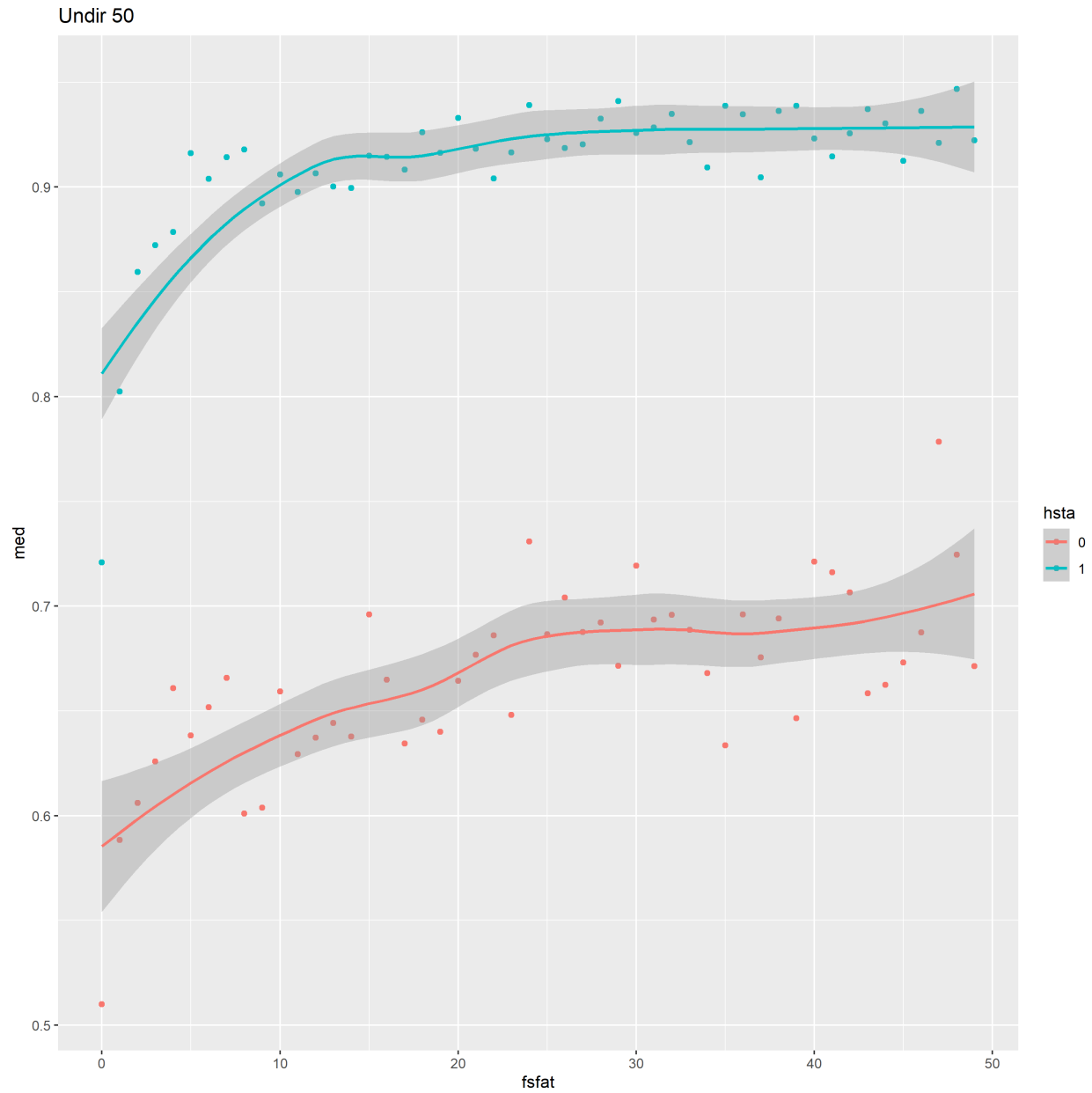


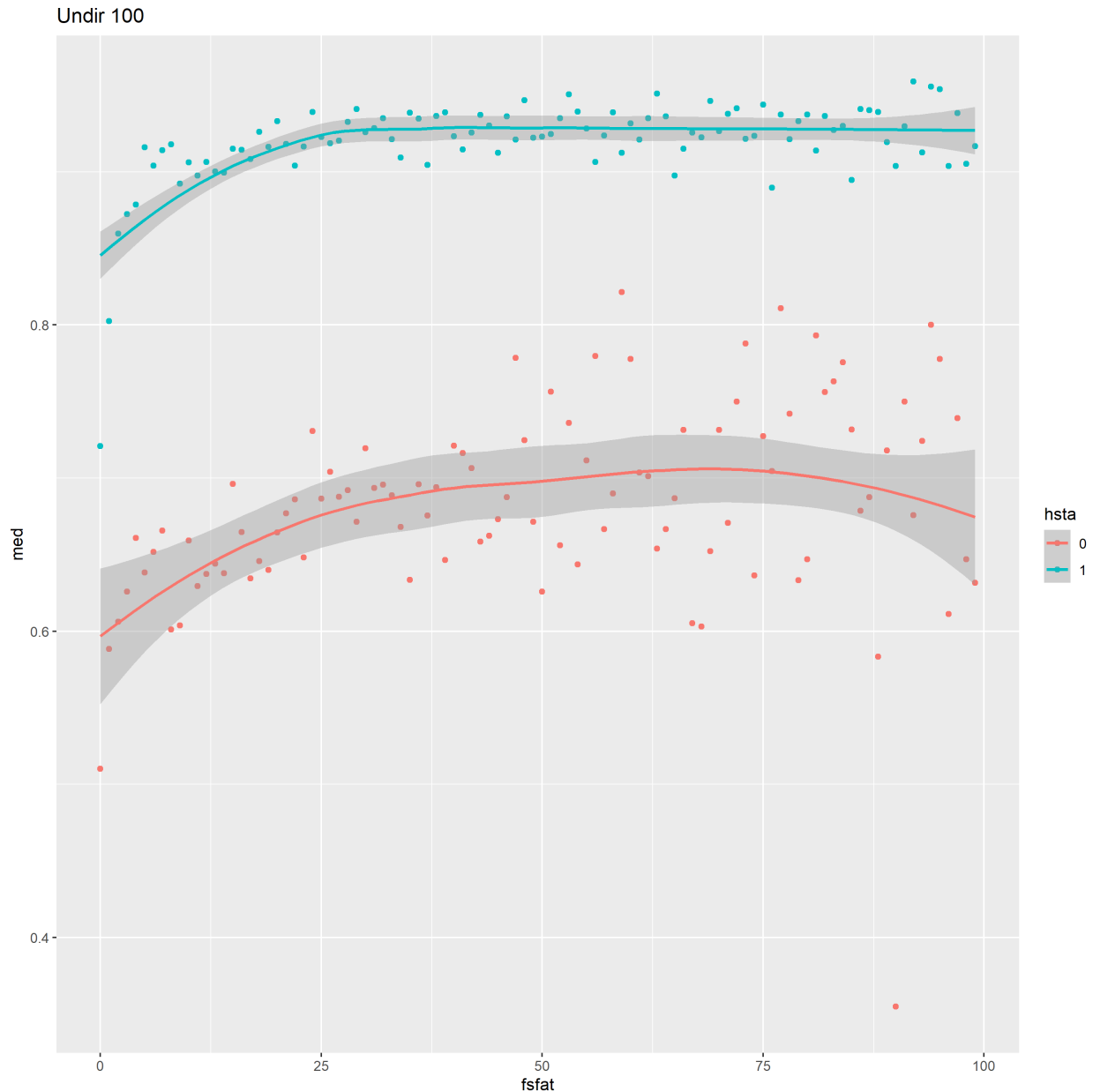
```
ggplot(hashLim100, aes(x = hsta)) +  
  geom_bar() +  
  ggtitle("undir 100")
```



Það sést að ef við tökum bara undir 50, þá er hsta orðið miklu nær og þegar við erum með undir 100, þá erum við kominn nálægt ójöfnunni eins og áðan.

Gott að skoða meðaltals myndina aftur, eftir að skera eftir limit.





Nú sést aðeins betur hvað er að gerast, í fyrri sjáum við að það lítur út fyrir að vera árangur og framfarir. Þar sem báðar línur eru að vaxa þegar fleiri spurninga hefur verið svarað. En það sést á seinni myndinni að það fer að jafna sig eftir það og framfarir hætta. Svo er það eitthvað smá fall hjá nýjum svörum.

3.3 umræður

Þá kemur spurningin hvort þetta er nokkuð framfarir eða utanaðbókarlærdómur, í fyrri þá lítur út fyrir að líkurnar vaxa. En það virðist svo jafna sig eftir það.

1. Ein hugmynd um hvað er að gerast þar er að þetta er ekki framfarir sem er að vaxa, heldur er það utanaðbókarlærdómur á röngu svarmöguleikunum, Það gæti verið aðal áhrifinn þar. Hugmynd til að skoða það betur væri kannski að skoða eins mynd, nema filterað bara fyrir spurningarnar sem hafa rétt svör sem eru að sjást í fyrsta skiptið og gert flokka fyrir hlutfall rangra svara sem hafa sést áður. Þá

mun koma ein lína fyrir hvert hlutfall, ef það er allt beinar láréttar línur, þá gæti það bent til að verið sé að ræða utanaðbókarlærdómur sem aðal áhrifin.

- Önnur hugmynd er að ef við gætum skipt nemendum upp í flokka, sem nemendur sem setja allt á minnið og nemendur sem eru að læra og fá framfarir. Svo er spurning hvort framfarirnar sjást betur þa.
- Loka hugmynd er að í staðinn fyrir að skoða fsfat eftir fyrirlestri, að skoða það kannski frekar eftir spurningahaus, þá gæti sést betur áhrif á sambærilegum spurningum eftir því hve margar tilraunir nemandinn hefur fengið á spurninguna sjálfa.

4 Modelgerð

Hér er smá grunnur af því sem ég er kominn með fyrir modelgerðina.

Ég byrjaði að skoða einföld glm, en þar sem það er verið að ræða um mixed effect model, þá hef ég fært yfir í glmer í staðinn. Ég ákvað að keyra allt miðað við 100 limitinn, ég hef ekki neina sterka ástæðu enn, þar sem ég veit ekki um leið til að sýna hvort einn gagnasett er betri en hitt, svo mér fannst 100 vera nægilega gott í bili.

4.1 gagnavinnsla

Fyrsta sem þarf að gera, er smá gagnavinnsla. Glmer tekur ekki vel við háu tölunum í fsfat, svo ég deili með 10 til að fá glmer til að keyra almennilega

```
hashLim100$fsfat <- hashLim100$fsfat/10
```

4.2 modelin sjálf

Modelinn sem ég er að vinna með á þessum tímapunkti er

- `glmer(correct ~ fsfat*hsta + nicc + gpow + lectureId + (1 | studentId), family = binomial(link = "logit"), ...)`
- `glmer(correct ~ fsfat*hsta + nicc + gpow + lectureId + (1 + fsfat * hsta | studentId), family = binomial(link = "logit"), ...)`

Stóri munurinn milli þeirra er að ég hef meira í mixed effect modelinu seinna en í því fyrra

Ég keyri módelinn ekki hér, því þau eru of löng, en hef save-að þau til seinni notkunar, svo hægt sé að load-a þeim upp

```
# #Þetta er fyrri modelið
# load("Data/ans22")
# #Þetta er seinna modelið
# load("Data/ans42")
```

4.2.1 Uppfært aðferð til að keyra hraðar

```
options(contrasts = c("contr.sum", "contr.poly"))
ans22 <- glmer(correct ~ fsfat*hsta + nicc + gpow + lectureId + (1 | studentId), family = binomial(link
  data = hashLim100, nAGQ = 0, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun
ans42 <- glmer(correct ~ fsfat*hsta + nicc + gpow + lectureId + (1 + fsfat * hsta | studentId),
  family = binomial(link = "logit"), data = hashLim100, nAGQ = 0,
  control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=2e5)))
```

4.3 smá model Analysis

Summaríð þeirra er svo

```
summary(ans22)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
## Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
## Family: binomial ( logit )
## Formula: correct ~ fsfat * hsta + nicc + gpow + lectureId + (1 | studentId)
## Data: hashLim100
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
## 62028.4 62253.4 -30990.2 61980.4    87232
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -11.9180   0.1437   0.2316   0.4062   3.7665
##
## Random effects:
##  Groups      Name      Variance Std.Dev.
## studentId (Intercept) 0.7562   0.8696
## Number of obs: 87256, groups: studentId, 283
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.659542   0.062155  26.700 < 2e-16 ***
## fsfat        0.203916   0.005618  36.297 < 2e-16 ***
## hsta1       -1.069679   0.019517 -54.808 < 2e-16 ***
## nicc1        0.611590   0.136589   4.478 7.55e-06 ***
## nicc2        0.309886   0.054664   5.669 1.44e-08 ***
## nicc3       -0.140135   0.029050  -4.824 1.41e-06 ***
## nicc4       -0.032914   0.037133  -0.886 0.375404
## nicc5       -0.211427   0.035790  -5.907 3.47e-09 ***
## nicc6       -0.273325   0.035506  -7.698 1.38e-14 ***
## gpow         0.127291   0.035618   3.574 0.000352 ***
## lectureId1  -1.067109   0.027788 -38.401 < 2e-16 ***
## lectureId2   0.093693   0.038362   2.442 0.014593 *
## lectureId3   0.615713   0.049407  12.462 < 2e-16 ***
## lectureId4  -0.582781   0.030428 -19.153 < 2e-16 ***
## lectureId5  -0.206038   0.033502  -6.150 7.75e-10 ***
## lectureId6   0.107225   0.038847   2.760 0.005777 **
## lectureId7   0.266837   0.042952   6.212 5.22e-10 ***
## lectureId8   0.452446   0.051391   8.804 < 2e-16 ***
```

```
## lectureId9    0.333222    0.051628    6.454 1.09e-10 ***
## lectureId10 -0.116136    0.041688   -2.786 0.005339 **
## lectureId11  0.076908    0.044417    1.732 0.083359 .
## lectureId12  0.154198    0.044362    3.476 0.000509 ***
## fsfat:hsta1  0.046067    0.005156    8.935 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Correlation matrix not shown by default, as p = 23 > 12.
## Use print(x, correlation=TRUE) or
##     vcov(x)           if you need it
```

```
summary(ans42)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##   Family: binomial ( logit )
## Formula: correct ~ fsfat * hsta + nicc + gpow + lectureId + (1 + fsfat *
##   hsta | studentId)
##   Data: hashLim100
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
## 61585.5 61894.9 -30759.8 61519.5    87223
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -12.1835   0.1301   0.2263   0.4023   3.0975
##
## Random effects:
##   Groups      Name      Variance Std.Dev. Corr
## studentId (Intercept) 0.661245 0.81317
##           fsfat       0.020476 0.14309   0.07
##           hsta1       0.066818 0.25849  -0.24  0.03
##           fsfat:hsta1 0.003598 0.05998   0.49  0.10 -0.63
## Number of obs: 87256, groups:  studentId, 283
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.571471   0.060301  26.060 < 2e-16 ***
## fsfat        0.257279   0.011425  22.520 < 2e-16 ***
## hsta1       -0.996831   0.026795 -37.202 < 2e-16 ***
## nicc1        0.616526   0.136683   4.511 6.46e-06 ***
## nicc2        0.315650   0.054950   5.744 9.23e-09 ***
## nicc3       -0.139659   0.029133  -4.794 1.64e-06 ***
## nicc4       -0.032797   0.037253  -0.880 0.37864
## nicc5       -0.215453   0.035925  -5.997 2.01e-09 ***
## nicc6       -0.275479   0.035656  -7.726 1.11e-14 ***
## gpow         0.123290   0.037942   3.249 0.00116 **
## lectureId1  -1.074036   0.028469 -37.726 < 2e-16 ***
## lectureId2   0.085378   0.038623   2.211 0.02707 *
## lectureId3   0.621246   0.049472  12.558 < 2e-16 ***
```

```
## lectureId4 -0.604654 0.030955 -19.533 < 2e-16 ***
## lectureId5 -0.203030 0.033852 -5.998 2.00e-09 ***
## lectureId6 0.110226 0.039069 2.821 0.00478 **
## lectureId7 0.275982 0.043225 6.385 1.72e-10 ***
## lectureId8 0.454300 0.051527 8.817 < 2e-16 ***
## lectureId9 0.341082 0.051850 6.578 4.76e-11 ***
## lectureId10 -0.108184 0.042085 -2.571 0.01015 *
## lectureId11 0.078091 0.044570 1.752 0.07976 .
## lectureId12 0.156135 0.044570 3.503 0.00046 ***
## fsfat:hsta1 0.036528 0.007074 5.164 2.42e-07 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Correlation matrix not shown by default, as p = 23 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)          if you need it
```

Sést að það er einhver munur hér á milli við styrkleika breytanna. Sem passar frá muninum þeirra.

Ef við skoðum aðeins Anova hjá þeim báðum

```
Anova(ans22, type = 3)
```

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: correct
##              Chisq Df Pr(>Chisq)
## (Intercept)  712.894  1 < 2.2e-16 ***
## fsfat        1317.475  1 < 2.2e-16 ***
## hsta         3003.912  1 < 2.2e-16 ***
## nicc         146.443  6 < 2.2e-16 ***
## gpow         12.772  1 0.0003519 ***
## lectureId    1883.178 12 < 2.2e-16 ***
## fsfat:hsta    79.834  1 < 2.2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Anova(ans42, type = 3)
```

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: correct
##              Chisq Df Pr(>Chisq)
## (Intercept)  679.149  1 < 2.2e-16 ***
## fsfat        507.145  1 < 2.2e-16 ***
## hsta         1383.977  1 < 2.2e-16 ***
## nicc         149.801  6 < 2.2e-16 ***
## gpow         10.559  1 0.001157 **
## lectureId    1864.030 12 < 2.2e-16 ***
## fsfat:hsta    26.666  1 2.419e-07 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Þá lítur út fyrir að allar breytur eru marktækar hjá þeim.

4.4 Model comparison

Hvor er svo sterkari? Skoðum aðeins anova milli þeirra

```
anova(ans22, ans42)
```

```
## Data: hashLim100
## Models:
## ans22: correct ~ fsfat * hsta + nicc + gpow + lectureId + (1 | studentId)
## ans42: correct ~ fsfat * hsta + nicc + gpow + lectureId + (1 + fsfat *
## ans42:      hsta | studentId)
##      npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
## ans22   24 62028 62253 -30990    61980
## ans42   33 61586 61895 -30760    61520 460.88  9 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Lítur út fyrir að ans42 er betra, það getur passað þar sem það er meira leift. En vandinn er að seinni modelið tekur u.þ.b. tvöfalt lengri tíma að keyra í hvert sinn, svo þá kemur spurningin “hve mikið betri er seinni miðað við fyrri?”

Setjum upp fall fyrir BrierScore og reiknum svo það fyrir þau bæði

```
BrierScore <- function(modl, df ) {
  predicted <- predict(modl, type = "response")
  truth <- df$correct
  return(mean((predicted-truth)^2))
}
SBrierScore <- function(modl, df) {
  predicted <- predict(modl, type = "response")
  truth <- df$correct
  Bs <- mean((predicted-truth)^2)
  Bmax <- mean(predict(modl, type = "response")) * (1-mean(predict(modl, type = "response")))
  return(1 - Bs/Bmax)
}

br22 <- SBrierScore(ans22, hashLim100)
br42 <- SBrierScore(ans42, hashLim100)

br22

## [1] 0.2411666

br42

## [1] 0.2557395

br22-br42

## [1] -0.0145729
```

Nú er brierScore-ið hjá seinna minna, en bara um u.þ.b. -0.014573. Ég játa að ég veit ekki alveg enn hvað það þýðir milli styrkleikjanna þeirra, hvort það sé rosalega mikill munur eða bara smár munur.

Næst er að reikna AUC þeirra.

```
Auc22 <- AUC(predict(ans22, type = "response"), hashLim100$correct)
Auc42 <- AUC(predict(ans42, type = "response"), hashLim100$correct)
```

```
Auc22
```

```
## [1] 0.8331737
```

```
Auc42
```

```
## [1] 0.842699
```

```
Auc42 - Auc22
```

```
## [1] 0.009525255
```

Hér sést að AUC er stærra hjá seinna modelinu en fyrra, um 0.009525. Ef ég skil Auc rétt, þá er seinni modelið með betra AUC.

En er þetta betra Brier og betra AUC nokkuð nógu mikið betra til að leifa tvöfalt lengri keyrslutíma fyrir eitthvað sem tekur svona langan tíma? Er ekki ennþá viss.

5 Nýjar niðurstöður frá nýrri breytu

5.1 Nýja breytan

Ég fékk aðgang að hash fyrir röngu svörinn hjá gögnunum. Með þeim gögnum hef ég búið til breytu sem ég kalla hluta, sem stendur fyrir “hlutfall rangra svara sem hafa sést áður”, semsagt þegar nemandi er að svara spurningu i. og hefur séð 50% af röngu svar möguleikunum áður, þá er hluta= 0.5.

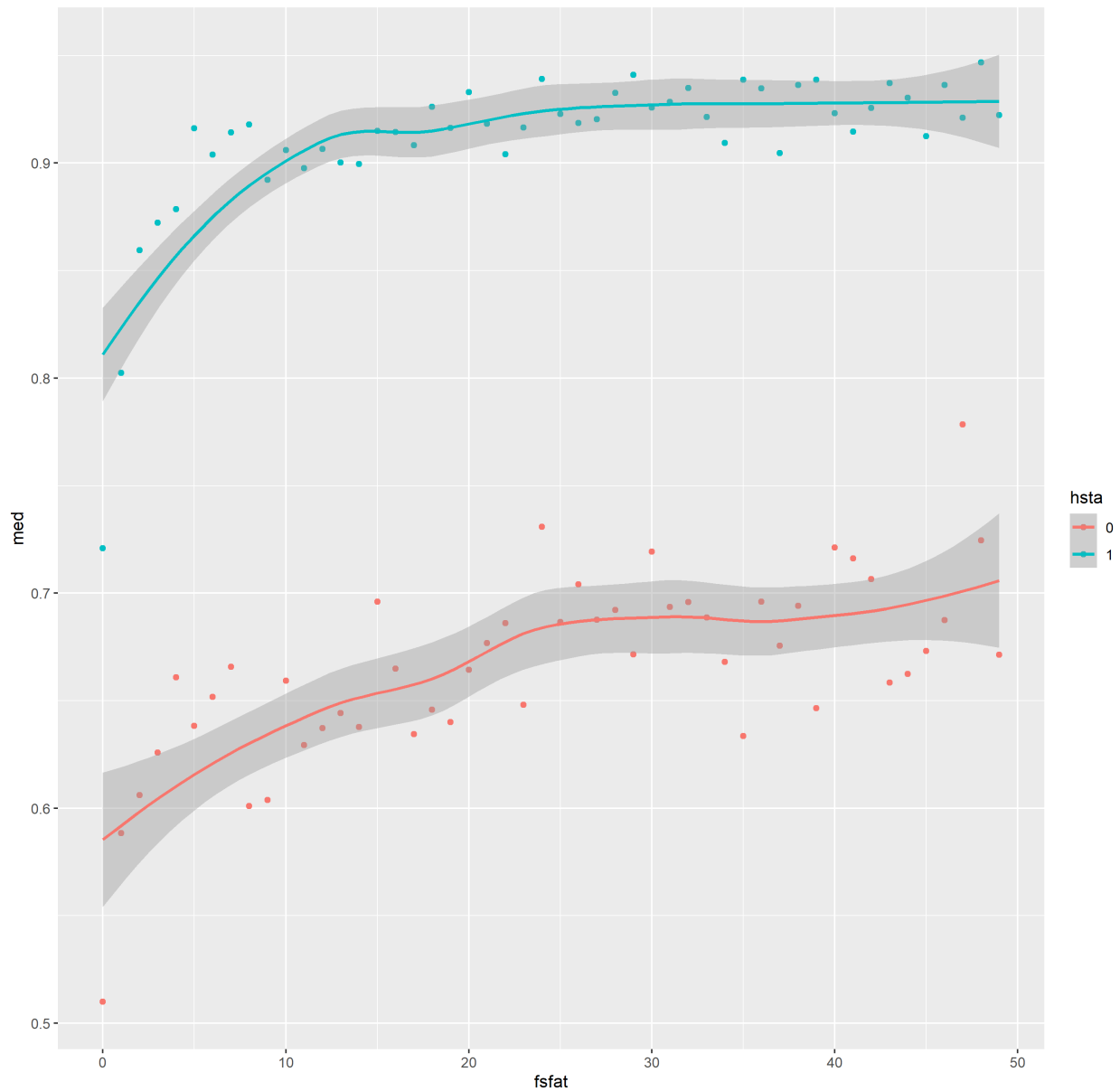
Það var samt smá hugsun fyrir aftan NOTA- og AOTA- spurningum, semsagt spurningar þar sem það er “none of the above” eða “All of the above” og það er vitlaust. Ég setti þau rök að ef NOTA- spurningu er að ræða og nemendann hefur séð rétta svarmöguleikann áður, þá er talið eins og NOTA- svarmöguleikinn hafi sést áður. Svo ef það er AOTA- spurning og nemandinn hefur séð eitthvað af hinum röngu svarmöguleikunum áður, þá segjum við að AOTA- svarið hafi sést áður.

Til að einfalda svo aðeins, fyrir teikningar. Þá bætti ég við hluta2 breytu, sem er discretized hluta breyta. Semsagt skip í 5 hluta.

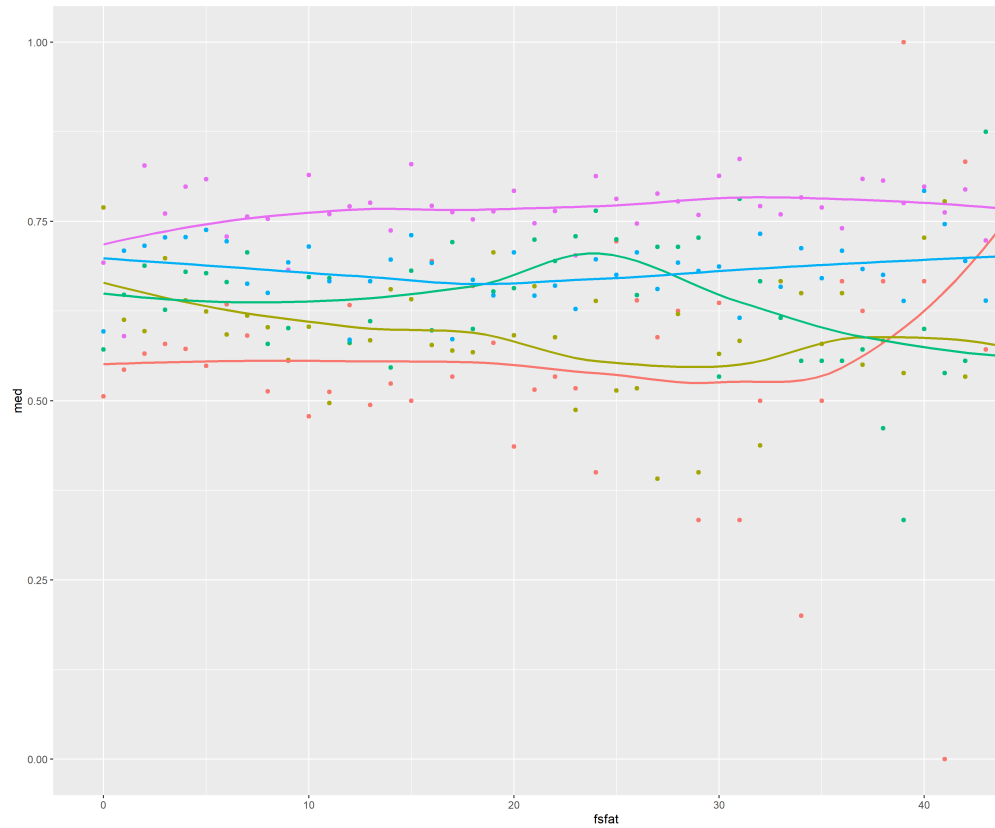
```
hashAnswer$hluta2 <- cut_interval(hashAnswer$hluta, n = 5)
hashLim50 <- hashAnswer %>% group_by(studentId) %>% mutate("count" = n()) %>%
  filter(count > 7 & fsfat < 50)
hashLim100 <- hashAnswer %>% group_by(studentId) %>% mutate("count" = n()) %>%
  filter(count > 7 & fsfat < 100)
hashLim50$fsfat <- hashLim50$fsfat/10
hashLim100$fsfat <- hashLim100$fsfat/10
```


5.2 Teikningar

Með þeirri nýju breytu er hægt að teikna nokkur gröf. Hér að ofan var teiknað meðaltals myndina Undir 50

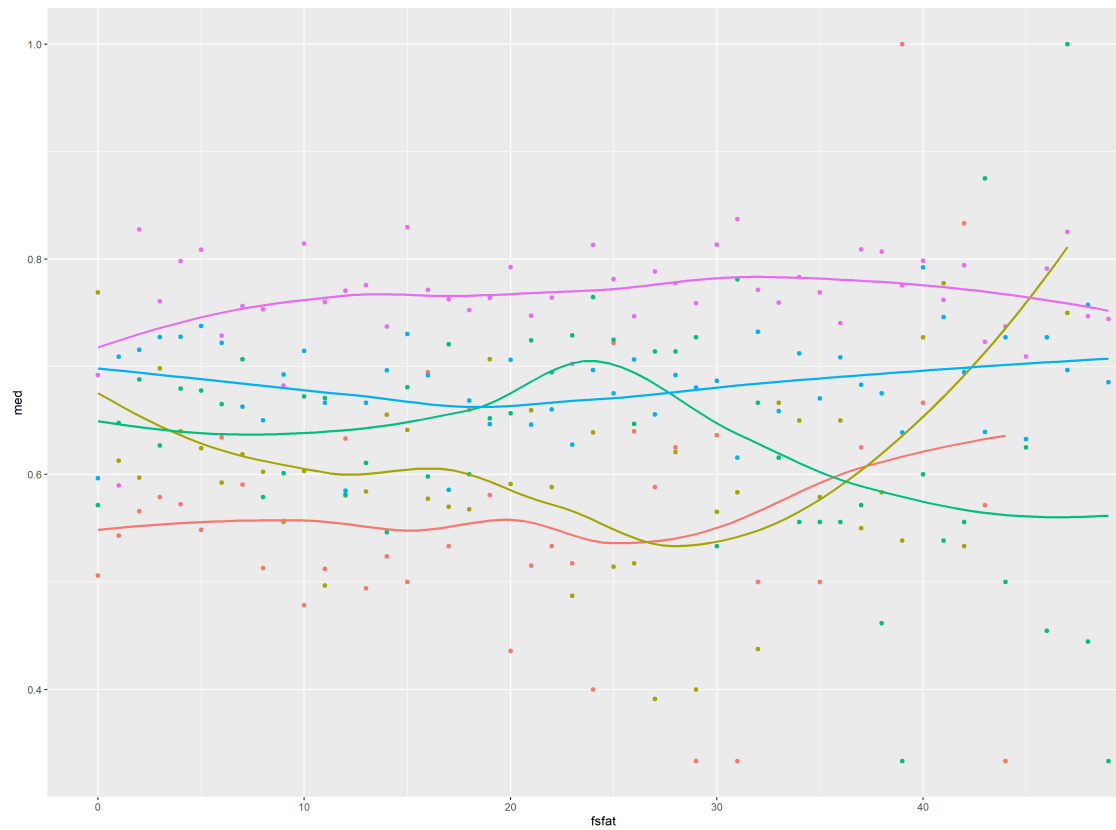


Þar sést að það er einhver vöxtur í gangi hjá spurningum með nýjum svörum á meðan fleiri spurningum er svarað. En hvernig líta meðaltals línurnar út ef flokkað er með tilliti til hluta2 flokkana, skoðað bara fyrir tilvikin



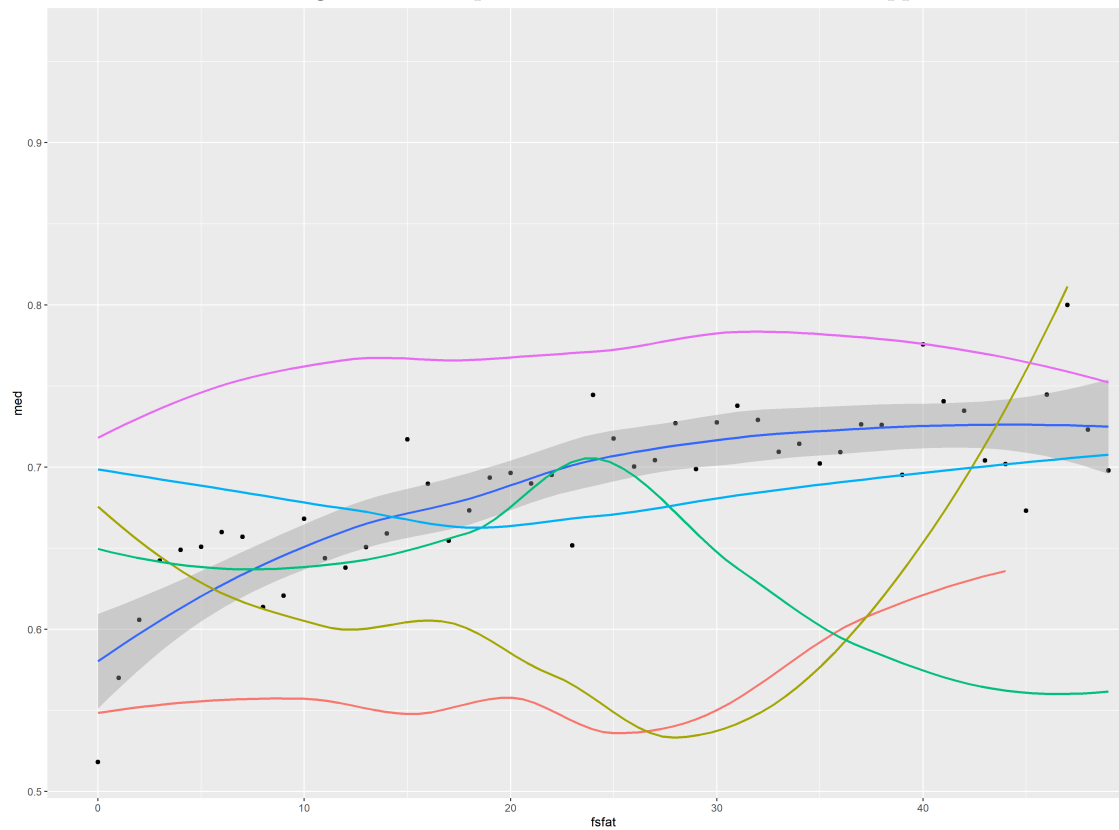
þegar $hsta == 0$? Þá fæst svona mynd:

Það er einhver stór vöxtur að koma hjá endanum, ég held að það kemur frá því að það eru ekki mikið af upplýsingum fyrir aftan suma punktana, svo ég teikna aftur upp myndina, nema að ég fjarlægji punkta þar sem ekki



eru fleiri en 5 gildi fyrir aftan

Hér sést að línurnar eru aðeins beinar hér. Gott viðmið gæti komið ef prufað er að teikna línurnar ofan á upp-



runalegu meðaltals myndina

Þetta bentir til að fyrir aftan vöxtinn sem sást áðan, þá eru einhverjar hlutfalls breytur að toga að aftan.

5.3 Töflur

Gott er að skoða aðeins fjöldana fyrir aftan töfluna hér að ofan T.d. er hægt að skoða hve mikið af gögnum það eru við hvert hlutfall

```
hashLim50 %>% filter(hsta == "0") %>% group_by(hluta2) %>%
  summarise(n())
```

```
## # A tibble: 5 x 2
##   hluta2   'n()'
##   <fct>   <int>
## 1 [0,0.2]   8285
## 2 (0.2,0.4] 4667
## 3 (0.4,0.6] 3868
## 4 (0.6,0.8] 7028
## 5 (0.8,1]   7226
```

5.4 uppfærð model

Þar sem komið er með nýja breytu, þá er hægt að skoða það miðað við upprunalega modelið og sjá hvernig það gengur Keyri þrjú mól

```
fit1 <- glmer(correct ~ fsfat*hsta + nicc + gpow + lectureId + (1 | studentId), family = binomial(link = logit),
  data = hashLim100, nAGQ = 0, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=10000))
fit3 <- glmer(correct ~ hluta2+hsta + nicc + gpow + lectureId + (1 | studentId), family = binomial(link = logit),
  data = hashLim100, nAGQ = 0, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=10000))
fit7 <- glmer(correct ~ hluta2 + fsfat + hsta + nicc + gpow + lectureId + (1 | studentId), family = binomial(link = logit),
  data = hashLim100, nAGQ = 0, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=10000))
```

Hér er: * fit1: upprunalega líkanið

- fit3: upprunalega líkanið, nema það er víxlað fsfat fyrir hluta2
- fit7: það er líkan sem hefur bæði fsfat og hluta2

Gott er kannski að skoða muninn á milli þeirra, með anova

```
anova(fit1, fit3)
```

```
## Data: hashLim100
## Models:
## fit1: correct ~ fsfat * hsta + nicc + gpow + lectureId + (1 | studentId)
## fit3: correct ~ hluta2 + hsta + nicc + gpow + lectureId + (1 | studentId)
##      npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## fit1    24 62028 62253 -30990    61980
## fit3    26 61571 61815 -30760    61519   461  2 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(fit3, fit7)
```

```
## Data: hashLim100
## Models:
## fit3: correct ~ hluta2 + hsta + nicc + gpow + lectureId + (1 | studentId)
## fit7: correct ~ hluta2 + fsfat + hsta + nicc + gpow + lectureId + (1 |
## fit7:      studentId)
##      npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
## fit3   26 61571 61815 -30760    61519
## fit7   27 61178 61431 -30562    61124 395.34   1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Hér segir að fit3 er marktækara en fit1 og fit7 er marktækara en fit3. Gott er að skoða aðeins standardized brier score muninn hjá þeim og svo líka AUC muninn

```
sb1 <- SBrierScore(fit1, hashLim100)
sb3 <- SBrierScore(fit3, hashLim100)
sb7 <- SBrierScore(fit7, hashLim100)

fitauc1 <- AUC(predict(fit1, type = "response"), hashLim100$correct)
fitauc3 <- AUC(predict(fit3, type = "response"), hashLim100$correct)
fitauc7 <- AUC(predict(fit7, type = "response"), hashLim100$correct)

sb1
```

```
## [1] 0.2411666
```

```
sb3
```

```
## [1] 0.2479733
```

```
sb7
```

```
## [1] 0.2524708
```

```
fitauc1
```

```
## [1] 0.8331737
```

```
fitauc3
```

```
## [1] 0.8358489
```

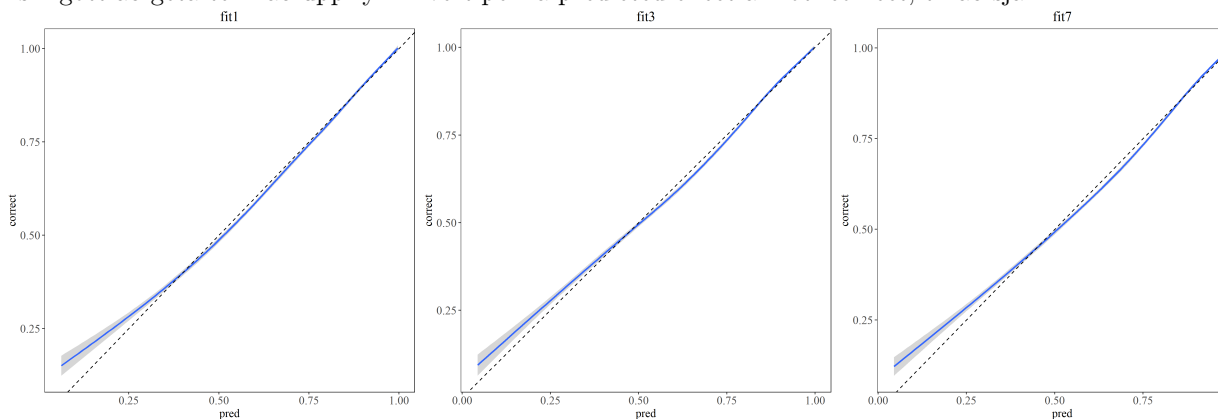
```
fitauc7
```

```
## [1] 0.8392987
```

Það er smár munur milli hverja, en ekki það mikill. Munurinn hjá standardized brier fyrir fit1 og fit3 er bara 0.0068. Sem segir til að munurinn á að nota hlutfall rangra svara séð í staðinn fyrir fjöldi spurninga fram að þessu er mjög lítill miðað við standardized brier. Það gæti verið að benda til að það gæti verið að áhrifinn að aftan vöxtinn hjá fsfat er aðallega bara hluta2. Að sama dúr er líka bara aðeins betra milli fit3 og fit7, með 0.0045, sem er enn minni en hinn. Svo á meðan það er marktækur munur á milli fit3 og fit7, þá er fit7 ekki endilega það mikið betra en fit3. Sömu niðurstöðurnar ef skoðað er munin hjá AUC í fit1 og fit3 0.0027 og munin milli fit3 og fit7 sem 0.0034. Á meðan AUC vex smá, þá er það ekki endilega nógu mikið til að segja til hvort eitt er betra en hitt. Ef ég skil þetta rétt.

5.5 Teikna fyrir model

Svo í lokinn, er kannski gott að geta teiknað upp fyrir hvert þeirra predicted effect á móti correct, til að sjá



hvernig það lítur út

Það lítur út fyrir að vera beint að mesta lagi, sem bendir að það er að ná vel á móti alvöru niðurstöðunum. En það er ekki mikill munur á milli þeirra. Svo eitt er ekki endilega betra en næsta?