A02g Gender Correlations with CRT control (P101 Final exam - Jan, 2018)

Joss Ives

June 7, 2018

*updated by Joss Ives 2018 June 07, 12:37:43*

# Overview

This is a companion report for “A02 Initial Explain Your Answee Report (P101 Final exam - Jan, 2018)”. In this report, correlations between the treatment and gender are examined in more detail. It has the same setup section as A02

# Setup

## Loading required package: Matrix

## Loading required package: xts

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

##   
## Attaching package: 'PerformanceAnalytics'

## The following object is masked from 'package:graphics':  
##   
## legend

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:xts':  
##   
## first, last

## The following objects are masked from 'package:plyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

### Add additional calculated values

The following is a summary of the variables present in the data file. The final 4 were caculated in this notebook. “Fix” refers to removing from the overall score on the test, the score of the specific question.

names(dat.raw)

## [1] "ID" "QNUM" "QCORRECT"   
## [4] "TREATMENT" "f.Atot40" "f.Btot38"   
## [7] "f.tot78" "course.grade" "d.version"   
## [10] "f.version" "NCRT" "Gender"   
## [13] "EYAfinal" "course.grade.frac" "CRT.medsplit"   
## [16] "final.grade.LMH" "final.grade.fix" "final.gradeA.fix"   
## [19] "f.tot100" "CONTROL" "GRELEVEL"

# Treatment x Gender

In this section we compare making data subset for each Gender with using the TREATMENT x Gender interaction terms

### Run the regressions on the two groups individually

Female-only data first

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ TREATMENT + final.grade.fix + NCRT + (1 | QNUM) +   
## (1 | ID)  
## Data: subset(dat.trt, Gender == "Female")  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 2021.1 2053.4 -1004.6 2009.1 1610   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.3893 -0.9279 0.4875 0.7773 2.7438   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.0000 0.0000   
## QNUM (Intercept) 0.1538 0.3921   
## Number of obs: 1616, groups: ID, 402; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.19254 0.29599 -7.407 1.29e-13 \*\*\*  
## TREATMENT1 0.27352 0.10748 2.545 0.0109 \*   
## final.grade.fix 2.95354 0.33856 8.724 < 2e-16 \*\*\*  
## NCRT 0.28429 0.05155 5.515 3.49e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) TREATM fnl.g.  
## TREATMENT1 -0.198   
## finl.grd.fx -0.644 0.020   
## NCRT -0.110 0.016 -0.256

## Est LL UL  
## (Intercept) 0.1116326 0.06249369 0.1994095  
## TREATMENT1 1.3145793 1.06487642 1.6228349  
## final.grade.fix 19.1736728 9.87446569 37.2303414  
## NCRT 1.3288131 1.20111718 1.4700849

Male-only data

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ TREATMENT + final.grade.fix + NCRT + (1 | QNUM) +   
## (1 | ID)  
## Data: subset(dat.trt, Gender == "Male")  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 1038.0 1067.4 -513.0 1026.0 978   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.8132 -0.7151 0.4209 0.5783 2.0031   
##   
## Random effects:  
## Groups Name Variance Std.Dev.   
## ID (Intercept) 1.607e-14 1.268e-07  
## QNUM (Intercept) 1.370e-01 3.701e-01  
## Number of obs: 984, groups: ID, 246; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.12431 0.37929 -5.601 2.13e-08 \*\*\*  
## TREATMENT1 0.15637 0.15477 1.010 0.312330   
## final.grade.fix 3.63811 0.48648 7.478 7.52e-14 \*\*\*  
## NCRT 0.27610 0.07786 3.546 0.000391 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) TREATM fnl.g.  
## TREATMENT1 -0.216   
## finl.grd.fx -0.727 0.021   
## NCRT -0.132 -0.003 -0.324

## Est LL UL  
## (Intercept) 0.1195155 0.0568290 0.2513495  
## TREATMENT1 1.1692592 0.8633132 1.5836281  
## final.grade.fix 38.0198543 14.6525039 98.6527170  
## NCRT 1.3179733 1.1314413 1.5352575

Gender effect for control condition only, controlling for question difficulty and test perfromance

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ Gender + final.grade.fix + NCRT + (1 | QNUM) + (1 |   
## ID)  
## Data: subset(dat.trt, TREATMENT == 0)  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 1546.5 1577.5 -767.2 1534.5 1294   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.1727 -0.8725 0.4335 0.7072 2.3491   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.01261 0.1123   
## QNUM (Intercept) 0.18857 0.4342   
## Number of obs: 1300, groups: ID, 648; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.32405 0.33678 -6.901 5.17e-12 \*\*\*  
## GenderMale 0.52005 0.13522 3.846 0.00012 \*\*\*  
## final.grade.fix 3.19719 0.39990 7.995 1.30e-15 \*\*\*  
## NCRT 0.27263 0.06148 4.434 9.25e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) GndrMl fnl.g.  
## GenderMale -0.047   
## finl.grd.fx -0.668 -0.069   
## NCRT -0.124 -0.085 -0.243

## Est LL UL  
## (Intercept) 0.09787614 0.05058332 0.1893853  
## GenderMale 1.68210471 1.29048754 2.1925638  
## final.grade.fix 24.46371898 11.17163691 53.5708018  
## NCRT 1.31340946 1.16429609 1.4816200

Gender effect for treatment condition only, controlling for question difficulty and test performance

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ Gender + final.grade.fix + NCRT + (1 | QNUM) + (1 |   
## ID)  
## Data: subset(dat.trt, TREATMENT == 1)  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 1516.7 1547.7 -752.3 1504.7 1294   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.3090 -0.9034 0.4542 0.6607 2.5177   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.06962 0.2639   
## QNUM (Intercept) 0.09734 0.3120   
## Number of obs: 1300, groups: ID, 648; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.07580 0.30212 -6.871 6.38e-12 \*\*\*  
## GenderMale 0.38585 0.13898 2.776 0.0055 \*\*   
## final.grade.fix 3.17993 0.40843 7.786 6.93e-15 \*\*\*  
## NCRT 0.29610 0.06345 4.667 3.06e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) GndrMl fnl.g.  
## GenderMale -0.038   
## finl.grd.fx -0.748 -0.081   
## NCRT -0.138 -0.101 -0.235

## Est LL UL  
## (Intercept) 0.1254565 0.06939482 0.2268085  
## GenderMale 1.4708620 1.12012317 1.9314260  
## final.grade.fix 24.0450052 10.79855733 53.5406961  
## NCRT 1.3446089 1.18737157 1.5226684

### Run the regressions on the entire data set with TREATMENT x Gender interaction term

First run with Gender=Female and TREATMENT=Control as the base levels

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ TREATMENT \* Gender + final.grade.fix + NCRT + (1 |   
## QNUM) + (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3047.2 3094.1 -1515.6 3031.2 2592   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.5634 -0.8913 0.4459 0.6890 2.7962   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.0000 0.0000   
## QNUM (Intercept) 0.1454 0.3813   
## Number of obs: 2600, groups: ID, 648; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.33411 0.26608 -8.772 < 2e-16 \*\*\*  
## TREATMENT1 0.27480 0.10784 2.548 0.010829 \*   
## GenderMale 0.50552 0.13203 3.829 0.000129 \*\*\*  
## final.grade.fix 3.18258 0.27773 11.459 < 2e-16 \*\*\*  
## NCRT 0.28318 0.04298 6.589 4.43e-11 \*\*\*  
## TREATMENT1:GenderMale -0.12384 0.18746 -0.661 0.508844   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) TREATMENT1 GndrMl fnl.g. NCRT   
## TREATMENT1 -0.214   
## GenderMale -0.099 0.401   
## finl.grd.fx -0.580 0.016 -0.067   
## NCRT -0.091 0.013 -0.075 -0.278   
## TREATMENT1: 0.118 -0.576 -0.693 0.000 -0.009

## Est LL UL  
## (Intercept) 0.09689658 0.05752008 0.163229  
## TREATMENT1 1.31626105 1.06548446 1.626061  
## GenderMale 1.65784027 1.27983308 2.147494  
## final.grade.fix 24.10880731 13.98850664 41.550868  
## NCRT 1.32734029 1.22011048 1.443994  
## TREATMENT1:GenderMale 0.88351783 0.61185061 1.275808

Then run with Gender=Male and TREATMENT=Control as the base levels

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ TREATMENT \* GRELEVEL + final.grade.fix + NCRT + (1 |   
## QNUM) + (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3047.2 3094.1 -1515.6 3031.2 2592   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.5634 -0.8913 0.4459 0.6890 2.7962   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.0000 0.0000   
## QNUM (Intercept) 0.1454 0.3813   
## Number of obs: 2600, groups: ID, 648; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.82860 0.28509 -6.414 1.42e-10 \*\*\*  
## TREATMENT1 0.15095 0.15329 0.985 0.324750   
## GRELEVELFemale -0.50552 0.13204 -3.829 0.000129 \*\*\*  
## final.grade.fix 3.18258 0.27774 11.459 < 2e-16 \*\*\*  
## NCRT 0.28318 0.04298 6.589 4.43e-11 \*\*\*  
## TREATMENT1:GRELEVELFemale 0.12384 0.18747 0.661 0.508861   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) TREATMENT1 GRELEV fnl.g. NCRT   
## TREATMENT1 -0.268   
## GRELEVELFml -0.371 0.566   
## finl.grd.fx -0.573 0.011 0.067   
## NCRT -0.119 -0.001 0.075 -0.278   
## TREATMENT1: 0.211 -0.818 -0.693 0.000 0.009

## Est LL UL  
## (Intercept) 0.1606391 0.09187044 0.2808837  
## TREATMENT1 1.1629403 0.86114072 1.5705100  
## GRELEVELFemale 0.6031946 0.46565635 0.7813567  
## final.grade.fix 24.1087681 13.98800641 41.5522185  
## NCRT 1.3273404 1.22011016 1.4439946  
## TREATMENT1:GRELEVELFemale 1.1318385 0.78380642 1.6344067

Then run with Gender=Male and TREATMENT=Treatment as the base levels

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ CONTROL \* GRELEVEL + final.grade.fix + NCRT + (1 |   
## QNUM) + (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3047.2 3094.1 -1515.6 3031.2 2592   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.5634 -0.8913 0.4459 0.6890 2.7962   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.0000 0.0000   
## QNUM (Intercept) 0.1454 0.3813   
## Number of obs: 2600, groups: ID, 648; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.67764 0.28517 -5.883 4.03e-09 \*\*\*  
## CONTROL1 -0.15095 0.15329 -0.985 0.32474   
## GRELEVELFemale -0.38167 0.13511 -2.825 0.00473 \*\*   
## final.grade.fix 3.18257 0.27774 11.459 < 2e-16 \*\*\*  
## NCRT 0.28318 0.04298 6.589 4.43e-11 \*\*\*  
## CONTROL1:GRELEVELFemale -0.12384 0.18746 -0.661 0.50885   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) CONTROL1 GRELEV fnl.g. NCRT   
## CONTROL1 -0.269   
## GRELEVELFml -0.382 0.582   
## finl.grd.fx -0.566 -0.011 0.066   
## NCRT -0.120 0.001 0.085 -0.278   
## CONTROL1:GR 0.228 -0.818 -0.710 0.000 -0.009

## Est LL UL  
## (Intercept) 0.1868139 0.1068244 0.3266990  
## CONTROL1 0.8598895 0.6367388 1.1612454  
## GRELEVELFemale 0.6827188 0.5238805 0.8897161  
## final.grade.fix 24.1087448 13.9880363 41.5520492  
## NCRT 1.3273403 1.2201103 1.4439943  
## CONTROL1:GRELEVELFemale 0.8835182 0.6118472 1.2758159

And then for completion run with Gender=Female and TREATMENT=Treatment as the base levels

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula:   
## QCORRECT ~ CONTROL \* Gender + final.grade.fix + NCRT + (1 | QNUM) +   
## (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3047.2 3094.1 -1515.6 3031.2 2592   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.5634 -0.8913 0.4459 0.6890 2.7962   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.0000 0.0000   
## QNUM (Intercept) 0.1454 0.3813   
## Number of obs: 2600, groups: ID, 648; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.05932 0.26486 -7.775 7.54e-15 \*\*\*  
## CONTROL1 -0.27479 0.10784 -2.548 0.01083 \*   
## GenderMale 0.38167 0.13511 2.825 0.00473 \*\*   
## final.grade.fix 3.18257 0.27774 11.459 < 2e-16 \*\*\*  
## NCRT 0.28318 0.04298 6.589 4.43e-11 \*\*\*  
## CONTROL1:GenderMale 0.12384 0.18747 0.661 0.50886   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) CONTROL1 GndrMl fnl.g. NCRT   
## CONTROL1 -0.192   
## GenderMale -0.099 0.407   
## finl.grd.fx -0.576 -0.016 -0.066   
## NCRT -0.086 -0.013 -0.085 -0.278   
## CONTROL1:GM 0.116 -0.576 -0.710 0.000 0.009

## Est LL UL  
## (Intercept) 0.1275413 0.07589241 0.2143400  
## CONTROL1 0.7597279 0.61498029 0.9385447  
## GenderMale 1.4647313 1.12395036 1.9088367  
## final.grade.fix 24.1087465 13.98798815 41.5521985  
## NCRT 1.3273405 1.22011041 1.4439945  
## CONTROL1:GenderMale 1.1318393 0.78380720 1.6344074

### Summary of results

*Format table in markdown instead of R for fun*

None of the cross terms were statistically significant

|  |  |  |
| --- | --- | --- |
| ODDS RATIOS [95%CI] | Data Subset | Interaction Term |
| Treatment (Female) | 1.315\* [1.065, 1.623] | 1.316\* [1.065, 1.626] |
| Treatment (Male) | 1.169 [0.863, 1.584] | 1.163 [0.861, 1.571] |
| Gender (Control) | 1.682\*\*\* [1.29, 2.193] | 1.658\*\*\* [1.28, 2.147] |
| Gender (Treatment) | 1.471\*\* [1.12, 1.931] | 1.465\*\* [1.124, 1.909] |