Analyses for Gonzalez et al. (2024)

Daniel K. Feinberg

Aug 27, 2024

Contents

# 1. Evaluate variables

## 1.1 Variable types

str(ordinalRegres\_substanceUse)

tibble [36,485 × 24] (S3: tbl\_df/tbl/data.frame)  
 $ region : chr [1:36485] "Arecibo" "Arecibo" "Arecibo" "Arecibo" ...  
 $ district : chr [1:36485] "Arecibo" "Arecibo" "Arecibo" "Arecibo" ...  
 $ grade : num [1:36485] 8 8 8 8 8 8 7 7 7 8 ...  
 $ adult\_ss : Factor w/ 2 levels "No","Yes": 2 2 2 2 2 2 2 2 2 2 ...  
 $ peer\_ss : Factor w/ 2 levels "No","Yes": 2 2 2 2 2 2 2 1 2 2 ...  
 $ teacher\_ss : Factor w/ 2 levels "No","Yes": 2 2 1 2 1 2 1 1 2 2 ...  
 $ s1 : num [1:36485] 0 1 0 0 0 0 0 0 0 0 ...  
 $ s2 : num [1:36485] 0 0 0 0 0 0 0 1 0 0 ...  
 $ s3 : num [1:36485] 2 0 1 0 0 0 4 0 1 1 ...  
 $ s4 : num [1:36485] 0 0 2 0 0 0 0 0 0 0 ...  
 $ s5 : num [1:36485] 0 0 0 0 0 0 0 0 2 0 ...  
 $ s6 : num [1:36485] 0 1 2 0 0 0 0 0 0 0 ...  
 $ s7 : num [1:36485] 0 0 0 0 0 0 0 0 0 0 ...  
 $ s8 : num [1:36485] 1 0 1 0 0 0 0 0 0 0 ...  
 $ s9 : num [1:36485] 2 0 0 0 0 0 1 0 2 0 ...  
 $ s10 : num [1:36485] 0 0 2 0 0 0 0 0 0 0 ...  
 $ s11 : num [1:36485] 0 0 2 0 0 0 1 0 0 0 ...  
 $ s12 : num [1:36485] 3 0 2 4 1 4 2 1 0 3 ...  
 $ substanceUse : Ord.factor w/ 3 levels "None"<"Low"<"High": 1 1 1 1 1 1 1 1 1 1 ...  
 $ gender : Factor w/ 2 levels "Male","Female": 2 2 2 2 2 2 2 2 2 2 ...  
 $ id : chr [1:36485] "000001" "000002" "000003" "000004" ...  
 $ has\_missing : logi [1:36485] FALSE FALSE FALSE FALSE FALSE FALSE ...  
 $ ptsd\_total : num [1:36485] 8 2 12 4 1 4 8 2 5 4 ...  
 $ social\_support\_total: Ord.factor w/ 4 levels "0 of 3 Endorsed"<..: 4 4 3 4 3 4 3 2 4 4 ...

## 1.2 Variable summaries

summary(ordinalRegres\_substanceUse)

region district grade adult\_ss peer\_ss   
 Length:36485 Length:36485 Min. : 7.000 No : 1859 No : 4570   
 Class :character Class :character 1st Qu.: 8.000 Yes:34626 Yes:31915   
 Mode :character Mode :character Median : 9.000   
 Mean : 9.383   
 3rd Qu.:11.000   
 Max. :12.000   
 teacher\_ss s1 s2 s3 s4   
 No :15964 Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.000   
 Yes:20521 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.000   
 Median :0.0000 Median :0.0000 Median :1.0000 Median :0.000   
 Mean :0.3649 Mean :0.4652 Mean :0.9974 Mean :0.577   
 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.000   
 Max. :4.0000 Max. :4.0000 Max. :4.0000 Max. :4.000   
 s5 s6 s7 s8   
 Min. :0.0000 Min. :0.00 Min. :0.0000 Min. :0.0000   
 1st Qu.:0.0000 1st Qu.:0.00 1st Qu.:0.0000 1st Qu.:0.0000   
 Median :0.0000 Median :0.00 Median :0.0000 Median :0.0000   
 Mean :0.4934 Mean :0.52 Mean :0.2277 Mean :0.7088   
 3rd Qu.:1.0000 3rd Qu.:1.00 3rd Qu.:0.0000 3rd Qu.:1.0000   
 Max. :4.0000 Max. :4.00 Max. :4.0000 Max. :4.0000   
 s9 s10 s11 s12 substanceUse  
 Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.00 None:34075   
 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00 Low : 1741   
 Median :0.0000 Median :0.0000 Median :0.0000 Median :1.00 High: 669   
 Mean :0.7415 Mean :0.2858 Mean :0.5383 Mean :1.66   
 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:3.00   
 Max. :4.0000 Max. :4.0000 Max. :4.0000 Max. :4.00   
 gender id has\_missing ptsd\_total   
 Male :17969 Length:36485 Mode :logical Min. : 0.00   
 Female:18516 Class :character FALSE:36485 1st Qu.: 3.00   
 Mode :character Median : 6.00   
 Mean : 7.58   
 3rd Qu.:11.00   
 Max. :48.00   
 social\_support\_total  
 0 of 3 Endorsed: 767   
 1 of 3 Endorsed: 3201   
 2 of 3 Endorsed:13690   
 3 of 3 Endorsed:18827

## 1.3 Verify order of dependent variable

# Verifying dependent variable order  
unique(ordinalRegres\_substanceUse$substanceUse)

[1] None Low High  
Levels: None < Low < High

# 2. Verify cell sizes

## 2.1 By ptsd\_total

# Verifying cell sizes for ptsd  
xtabs(~ substanceUse + ptsd\_total,  
 data = ordinalRegres\_substanceUse  
)

ptsd\_total  
substanceUse 0 1 2 3 4 5 6 7 8 9 10 11 12  
 None 3166 2607 2634 3017 2851 2554 2279 1990 1790 1596 1337 1163 1013  
 Low 41 44 61 88 82 99 99 96 90 78 97 83 96  
 High 33 26 17 27 32 33 27 25 36 29 38 28 33  
 ptsd\_total  
substanceUse 13 14 15 16 17 18 19 20 21 22 23 24 25  
 None 852 740 655 574 496 434 378 288 274 241 197 163 131  
 Low 79 68 76 56 61 50 40 31 31 30 25 32 19  
 High 15 22 32 13 23 17 9 29 12 10 12 13 12  
 ptsd\_total  
substanceUse 26 27 28 29 30 31 32 33 34 35 36 37 38  
 None 110 109 79 73 51 42 38 43 30 13 13 13 11  
 Low 18 14 7 12 5 4 5 1 7 5 0 3 1  
 High 12 5 5 4 2 3 2 2 2 2 7 5 5  
 ptsd\_total  
substanceUse 39 40 41 42 43 44 45 46 47 48  
 None 8 4 4 5 3 1 1 1 2 1  
 Low 2 1 1 1 1 0 0 0 0 1  
 High 0 2 2 0 2 2 0 0 1 1

## 2.2 By social\_support\_total

# Verifying cell sizes for social support  
xtabs(~ substanceUse + social\_support\_total,  
 data = ordinalRegres\_substanceUse  
)

social\_support\_total  
substanceUse 0 of 3 Endorsed 1 of 3 Endorsed 2 of 3 Endorsed 3 of 3 Endorsed  
 None 668 2930 12667 17810  
 Low 54 189 742 756  
 High 45 82 281 261

## 2.3 By gender

# Verifying cell sizes for gender  
xtabs(~ substanceUse + gender,  
 data = ordinalRegres\_substanceUse  
)

gender  
substanceUse Male Female  
 None 16694 17381  
 Low 860 881  
 High 415 254

## 2.4 By grade

# Verifying cell sizes for grade  
xtabs(~ substanceUse + grade,  
 data = ordinalRegres\_substanceUse  
)

grade  
substanceUse 7 8 9 10 11 12  
 None 6609 6422 5500 5476 5468 4600  
 Low 104 146 222 304 415 550  
 High 41 52 85 106 160 225

# 3. Cronbach’s alpha

c\_alpha <- ordinalRegres\_substanceUse %>%  
 dplyr::select(s1, s2, s3, s4, s5, s6, s7, s8, s9, s10, s11, s12)  
  
psych::alpha(c\_alpha)

Reliability analysis   
Call: psych::alpha(x = c\_alpha)  
  
 raw\_alpha std.alpha G6(smc) average\_r S/N ase mean sd median\_r  
 0.79 0.8 0.81 0.25 3.9 0.0016 0.63 0.56 0.21  
  
 95% confidence boundaries   
 lower alpha upper  
Feldt 0.78 0.79 0.79  
Duhachek 0.78 0.79 0.79  
  
 Reliability if an item is dropped:  
 raw\_alpha std.alpha G6(smc) average\_r S/N alpha se var.r med.r  
s1 0.77 0.78 0.79 0.25 3.6 0.0017 0.013 0.21  
s2 0.78 0.80 0.81 0.26 3.9 0.0017 0.013 0.21  
s3 0.78 0.79 0.80 0.25 3.7 0.0017 0.014 0.21  
s4 0.77 0.78 0.78 0.24 3.5 0.0018 0.012 0.21  
s5 0.77 0.78 0.79 0.24 3.5 0.0018 0.013 0.21  
s6 0.77 0.78 0.79 0.24 3.5 0.0017 0.012 0.21  
s7 0.77 0.78 0.79 0.24 3.5 0.0017 0.013 0.21  
s8 0.77 0.78 0.79 0.24 3.5 0.0018 0.012 0.21  
s9 0.76 0.77 0.78 0.24 3.4 0.0019 0.012 0.21  
s10 0.77 0.78 0.79 0.24 3.6 0.0017 0.013 0.21  
s11 0.77 0.78 0.79 0.25 3.6 0.0018 0.015 0.21  
s12 0.78 0.79 0.80 0.25 3.7 0.0017 0.015 0.21  
  
 Item statistics   
 n raw.r std.r r.cor r.drop mean sd  
s1 36485 0.51 0.54 0.48 0.40 0.36 0.84  
s2 36485 0.43 0.43 0.33 0.31 0.47 0.91  
s3 36485 0.55 0.52 0.45 0.41 1.00 1.18  
s4 36485 0.58 0.60 0.57 0.47 0.58 0.96  
s5 36485 0.58 0.57 0.52 0.47 0.49 0.98  
s6 36485 0.55 0.57 0.52 0.43 0.52 1.03  
s7 36485 0.54 0.59 0.55 0.46 0.23 0.68  
s8 36485 0.61 0.59 0.55 0.49 0.71 1.15  
s9 36485 0.66 0.63 0.61 0.55 0.74 1.16  
s10 36485 0.52 0.57 0.52 0.43 0.29 0.74  
s11 36485 0.57 0.56 0.49 0.45 0.54 1.03  
s12 36485 0.54 0.49 0.41 0.36 1.66 1.38  
  
Non missing response frequency for each item  
 0 1 2 3 4 miss  
s1 0.78 0.14 0.04 0.02 0.02 0  
s2 0.72 0.16 0.06 0.02 0.03 0  
s3 0.44 0.33 0.11 0.06 0.07 0  
s4 0.65 0.22 0.08 0.03 0.03 0  
s5 0.73 0.14 0.06 0.03 0.03 0  
s6 0.73 0.14 0.06 0.04 0.04 0  
s7 0.87 0.08 0.03 0.02 0.01 0  
s8 0.63 0.19 0.08 0.05 0.06 0  
s9 0.61 0.20 0.08 0.05 0.06 0  
s10 0.83 0.10 0.04 0.02 0.01 0  
s11 0.71 0.15 0.06 0.04 0.04 0  
s12 0.27 0.25 0.12 0.24 0.11 0

# 4. Ordinal (logit) multiple linear regression

## 4.1 Run the model

In this analysis, we conducted an ordinal (logit) multiple linear regression on the response variable self-reported substance use to see how it changes with respect to the following predictor variables:

1. social support from an adult at home *(yes/no)*
2. social support from at least one peer *(yes/no)*
3. social support from a teacher or counselor at school *(yes/no)*
4. ptsd\_total *(continuous)*
5. grade *(ordinal)*
6. gender *(categorical)*
7. gender and ptsd\_total *(interaction)*
8. gender and adult social support *(interaction)*
9. gender and peer social support *(interaction)*
10. gender and teacher/counselor social support *(interaction)*

Here’s our model:

clm\_lm <- clm(  
 substanceUse ~  
 adult\_ss +  
 peer\_ss +  
 teacher\_ss +  
 ptsd\_total +  
 grade +  
 gender +  
 gender \* ptsd\_total +  
 gender \* adult\_ss +  
 gender \* peer\_ss +  
 gender \* teacher\_ss,  
 data = ordinalRegres\_substanceUse,  
 link = "logit"  
)  
  
# Null model  
null\_clm <- clm(substanceUse ~ 1,  
 data = ordinalRegres\_substanceUse,  
 link = "logit"  
)  
  
# Full summary of ordinal regression  
clm\_full\_summary <- summary(clm\_lm)  
  
clm\_full\_summary # Return the model

formula:   
substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade + gender + gender \* ptsd\_total + gender \* adult\_ss + gender \* peer\_ss + gender \* teacher\_ss  
data: ordinalRegres\_substanceUse  
  
 link threshold nobs logLik AIC niter max.grad cond.H   
 logit flexible 36485 -9287.90 18599.80 7(0) 2.28e-11 6.2e+04  
  
Coefficients:  
 Estimate Std. Error z value Pr(>|z|)   
adult\_ssYes -0.864663 0.107521 -8.042 8.85e-16 \*\*\*  
peer\_ssYes 0.273027 0.091584 2.981 0.002872 \*\*   
teacher\_ssYes -0.241849 0.062463 -3.872 0.000108 \*\*\*  
ptsd\_total 0.086179 0.003869 22.276 < 2e-16 \*\*\*  
grade 0.387776 0.014076 27.549 < 2e-16 \*\*\*  
genderFemale -0.195315 0.178714 -1.093 0.274442   
ptsd\_total:genderFemale -0.014565 0.005342 -2.726 0.006403 \*\*   
adult\_ssYes:genderFemale -0.023452 0.151757 -0.155 0.877185   
peer\_ssYes:genderFemale -0.034089 0.140895 -0.242 0.808823   
teacher\_ssYes:genderFemale -0.049490 0.090310 -0.548 0.583690   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Threshold coefficients:  
 Estimate Std. Error z value  
None|Low 6.3857 0.1824 35.01  
Low|High 7.7850 0.1862 41.82

## 4.2 Calculate confidence intervals

# Extract coefficients from the summary  
clm\_coefs <- clm\_full\_summary$coefficients  
  
# Calculate confidence intervals from ordinal regression  
clm\_confints\_pred <- confint(clm\_lm, level = 0.95)  
  
# Extract the coefficients and variance-covariance matrix  
isolated\_clm\_coefs <- coef(clm\_lm)  
clm\_vcov <- vcov(clm\_lm)  
  
# Compute standard error using variance-covariance matrix  
clm\_se <- sqrt(diag(clm\_vcov))  
  
# Compute the 95% confidence interval and add to clm\_confints  
all\_clm\_coef\_ci <- cbind(  
 "estimate" = isolated\_clm\_coefs,  
 "2.5 %" = isolated\_clm\_coefs - 1.96 \* clm\_se,  
 "97.5 %" = isolated\_clm\_coefs + 1.96 \* clm\_se  
)  
  
# Use cbind() to add the new coefficients column  
clm\_coefs <- cbind(  
 all\_clm\_coef\_ci,  
 clm\_coefs  
)  
  
# Remove prior estimates column  
clm\_coefs <- clm\_coefs[, -4]  
  
# Return calculated confidence intervals  
clm\_coefs

estimate 2.5 % 97.5 % Std. Error  
None|Low 6.38565726 6.02819317 6.743121345 0.182379636  
Low|High 7.78496119 7.42010130 8.149821084 0.186153005  
adult\_ssYes -0.86466262 -1.07540289 -0.653922344 0.107520548  
peer\_ssYes 0.27302656 0.09352193 0.452531182 0.091583992  
teacher\_ssYes -0.24184859 -0.36427599 -0.119421191 0.062462960  
ptsd\_total 0.08617895 0.07859647 0.093761420 0.003868609  
grade 0.38777580 0.36018733 0.415364262 0.014075748  
genderFemale -0.19531511 -0.54559500 0.154964786 0.178714231  
ptsd\_total:genderFemale -0.01456537 -0.02503632 -0.004094419 0.005342321  
adult\_ssYes:genderFemale -0.02345227 -0.32089526 0.273990724 0.151756629  
peer\_ssYes:genderFemale -0.03408866 -0.31024218 0.242064868 0.140894656  
teacher\_ssYes:genderFemale -0.04949000 -0.22649695 0.127516942 0.090309666  
 z value Pr(>|z|)  
None|Low 35.0129949 1.427000e-268  
Low|High 41.8202283 0.000000e+00  
adult\_ssYes -8.0418361 8.850225e-16  
peer\_ssYes 2.9811603 2.871585e-03  
teacher\_ssYes -3.8718721 1.080026e-04  
ptsd\_total 22.2764660 6.249896e-110  
grade 27.5492142 4.522343e-167  
genderFemale -1.0928906 2.744418e-01  
ptsd\_total:genderFemale -2.7264120 6.402702e-03  
adult\_ssYes:genderFemale -0.1545387 8.771850e-01  
peer\_ssYes:genderFemale -0.2419443 8.088233e-01  
teacher\_ssYes:genderFemale -0.5480034 5.836896e-01

## 4.3 Evaluate model fit

### 4.3.1 Nagelkerke

nagelkerke(  
 fit = clm\_lm,  
 null = null\_clm  
)

$Models  
   
Model: "clm, substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade + gender + gender \* ptsd\_total + gender \* adult\_ss + gender \* peer\_ss + gender \* teacher\_ss, ordinalRegres\_substanceUse, logit"  
Null: "clm, substanceUse ~ 1, ordinalRegres\_substanceUse, logit"   
  
$Pseudo.R.squared.for.model.vs.null  
 Pseudo.R.squared  
McFadden 0.0983259  
Cox and Snell (ML) 0.0540071  
Nagelkerke (Cragg and Uhler) 0.1251780  
  
$Likelihood.ratio.test  
 Df.diff LogLik.diff Chisq p.value  
 -10 -1012.8 2025.7 0  
  
$Number.of.observations  
   
Model: 36485  
Null: 36485  
  
$Messages  
[1] "Note: For models fit with REML, these statistics are based on refitting with ML"  
  
$Warnings  
[1] "None"

### 4.3.2 Test parallel lines/proportional odds

#### 4.3.2.1 Brandt test

# Rerun model using polr() for brant package  
polr\_lm <- polr(  
 substanceUse ~  
 adult\_ss +  
 peer\_ss +  
 teacher\_ss +  
 ptsd\_total +  
 grade +  
 gender +  
 gender \* ptsd\_total +  
 gender \* adult\_ss +  
 gender \* peer\_ss +  
 gender \* teacher\_ss,  
 data = ordinalRegres\_substanceUse,  
 Hess = TRUE  
)  
  
# global brand test  
brant(polr\_lm)

------------------------------------------------------------   
Test for X2 df probability   
------------------------------------------------------------   
Omnibus 38.61 10 0  
adult\_ssYes 0.28 1 0.59  
peer\_ssYes 5.01 1 0.03  
teacher\_ssYes 0 1 0.98  
ptsd\_total 0.37 1 0.55  
grade 0.4 1 0.52  
genderFemale 2.17 1 0.14  
ptsd\_total:genderFemale 2.24 1 0.13  
adult\_ssYes:genderFemale 1.63 1 0.2  
peer\_ssYes:genderFemale 0.39 1 0.53  
teacher\_ssYes:genderFemale 0.02 1 0.89  
------------------------------------------------------------   
  
H0: Parallel Regression Assumption holds

### 4.3.3 Fit and significance for null and alternative models using broom::glance():

clm\_fit\_stats <- rbind(  
 "clm\_model" = broom::glance(clm\_lm),  
 "null\_model" = broom::glance(null\_clm)  
) %>%  
 clean\_names()  
  
# Return output:  
clm\_fit\_stats

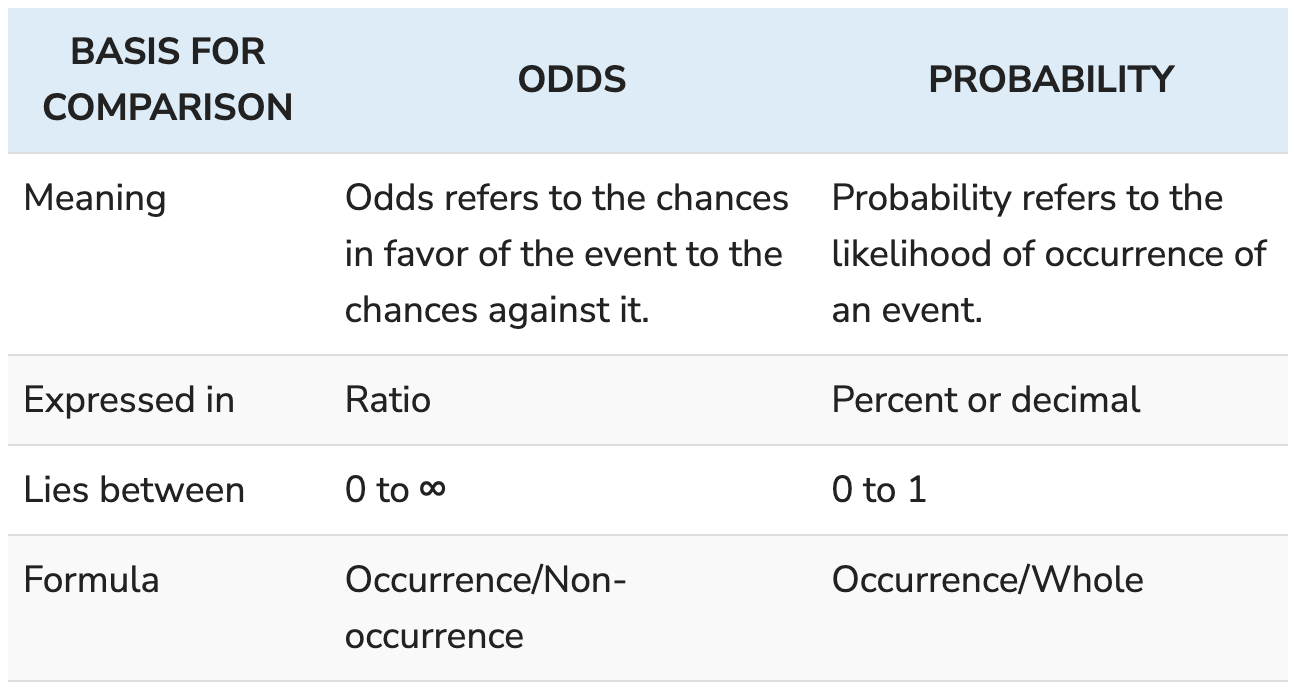
# A tibble: 2 × 6  
 edf aic bic log\_lik df\_residual nobs  
\* <int> <dbl> <dbl> <logLik> <dbl> <dbl>  
1 12 18600. 18702. -9287.899 36473 36485  
2 2 20605. 20622. -10300.728 36483 36485

## 4.4 Results table: log odds

clm\_tibble <- as\_tibble(clm\_coefs)  
clm\_tibble

# A tibble: 12 × 6  
 estimate `2.5 %` `97.5 %` `Std. Error` `z value` `Pr(>|z|)`  
 <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
 1 6.39 6.03 6.74 0.182 35.0 1.43e-268  
 2 7.78 7.42 8.15 0.186 41.8 0   
 3 -0.865 -1.08 -0.654 0.108 -8.04 8.85e- 16  
 4 0.273 0.0935 0.453 0.0916 2.98 2.87e- 3  
 5 -0.242 -0.364 -0.119 0.0625 -3.87 1.08e- 4  
 6 0.0862 0.0786 0.0938 0.00387 22.3 6.25e-110  
 7 0.388 0.360 0.415 0.0141 27.5 4.52e-167  
 8 -0.195 -0.546 0.155 0.179 -1.09 2.74e- 1  
 9 -0.0146 -0.0250 -0.00409 0.00534 -2.73 6.40e- 3  
10 -0.0235 -0.321 0.274 0.152 -0.155 8.77e- 1  
11 -0.0341 -0.310 0.242 0.141 -0.242 8.09e- 1  
12 -0.0495 -0.226 0.128 0.0903 -0.548 5.84e- 1

## 4.5 Reminder: probability vs. odds



[Citation](https://keydifferences.com/difference-between-odds-and-probability.html)

## 4.6 Results: odds ratios with interpretation

odds\_ratio\_clm <- exp(all\_clm\_coef\_ci)  
odds\_ratio\_clm

estimate 2.5 % 97.5 %  
None|Low 593.2745392 414.9645821 848.2041456  
Low|High 2404.1728472 1669.2025932 3462.7594654  
adult\_ssYes 0.4211936 0.3411603 0.5200021  
peer\_ssYes 1.3139351 1.0980347 1.5722869  
teacher\_ssYes 0.7851751 0.6946994 0.8874339  
ptsd\_total 1.0900014 1.0817677 1.0982977  
grade 1.4736993 1.4335979 1.5149225  
genderFemale 0.8225754 0.5794969 1.1676168  
ptsd\_total:genderFemale 0.9855402 0.9752745 0.9959140  
adult\_ssYes:genderFemale 0.9768206 0.7254992 1.3152026  
peer\_ssYes:genderFemale 0.9664858 0.7332693 1.2738768  
teacher\_ssYes:genderFemale 0.9517147 0.7973218 1.1360041

or\_tibble <- as\_tibble(odds\_ratio\_clm) # convert to tibble format for easy in-text reference

### 4.6.1 Significant predictor variables

* For students that endorsed having **social support from an adult**, the odds of endorsing high substance use (vs. low or no substance use) decrease by a factor of approximately 0.42.
* For students that endorsed having **social support from at least one peer**, the odds of endorsing high substance use (vs. low or no substance use) increased by a factor of approximately 1.31,
* For students that endorsed having **social support from a counselor or teacher at school**, the odds of endorsing high substance use (vs. low or no substance use) decreased by a factor of approximately 0.79.
* For each one unit increase in **PTSD Symptom Total**, the odds of a student endorsing high substance use (vs. low or no substance use) increased by a factor of approximately 1.09.
* For each increase in **Grade level**, the odds of endorsing high substance (vs. low or no substance use) use increased by a factor of approximately 1.47,
* Lastly, an interaction was evidenced, such that if a student was **female and experiencing ptsd symptoms**, the odds of endorsing high substance use (vs. low or no substance use) decreased by a factor of approximately 0.99 when compared to males experiencing ptsd symptoms.

### 4.6.2 Significant intercepts

* *Please note: model intercepts are generally not interpreted.*
* **No Substance Use Intercept:** This intercept coefficient value (593.27) represents the odds of a student’s self-endorsement of substance use having fallen within the *no substance use category* given values of 0 for all other coefficients (which is impossible).
* **Low Substance Use Intercept:** This intercept coefficient value (2404.17) represents the odds of a student’s self-endorsement of substance use having fallen within the *low substance use category* given values of 0 for all other coefficients (which is impossible).

## 4.7 Formatted regression table

### 4.7.1 JAMA Table for Regression (v1)

Table 1: Ordinal Regression Analysis of Outcome by Age and Gender

| Predictor | OR Estimates (95% CI)1 | Z-value | P-value |
| --- | --- | --- | --- |
| **adult\_ss** |  |  |  |
| No | — | — |  |
| Yes | 0.42 (0.34 to 0.52) | -8.04 | <0.001 |
| **peer\_ss** |  |  |  |
| No | — | — |  |
| Yes | 1.31 (1.10 to 1.58) | 2.98 | 0.003 |
| **teacher\_ss** |  |  |  |
| No | — | — |  |
| Yes | 0.79 (0.69 to 0.89) | -3.87 | <0.001 |
| **ptsd\_total** | 1.09 (1.08 to 1.10) | 22.3 | <0.001 |
| **Grade** | 1.47 (1.43 to 1.52) | 27.5 | <0.001 |
| **gender** |  |  |  |
| Male | — | — |  |
| Female | 0.82 (0.58 to 1.17) | -1.09 | 0.27 |
| **ptsd\_total \* gender** |  |  |  |
| ptsd\_total \* Female | 0.99 (0.98 to 1.00) | -2.73 | 0.006 |
| **adult\_ss \* gender** |  |  |  |
| Yes \* Female | 0.98 (0.73 to 1.32) | -0.155 | 0.88 |
| **peer\_ss \* gender** |  |  |  |
| Yes \* Female | 0.97 (0.73 to 1.28) | -0.242 | 0.81 |
| **teacher\_ss \* gender** |  |  |  |
| Yes \* Female | 0.95 (0.80 to 1.14) | -0.548 | 0.58 |
| 1OR = Odds Ratio, CI = Confidence Interval | | | |

# 5. Exploring interaction effect of ptsd\_total:gender\_female

## 5.1 ANOVA method to determine how interaction affects model fit

Likelihood ratio tests of cumulative link models:  
   
 formula:   
clm\_lm\_alt substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade + gender + gender \* adult\_ss + gender \* peer\_ss + gender \* teacher\_ss   
clm\_lm substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade + gender + gender \* ptsd\_total + gender \* adult\_ss + gender \* peer\_ss + gender \* teacher\_ss  
 link: threshold:  
clm\_lm\_alt logit flexible   
clm\_lm logit flexible   
  
 no.par AIC logLik LR.stat df Pr(>Chisq)   
clm\_lm\_alt 11 18605 -9291.6   
clm\_lm 12 18600 -9287.9 7.4317 1 0.006409 \*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## 5.2 Compare models: no interaction vs significant interaction

Likelihood ratio tests of cumulative link models:  
   
 formula:   
clm\_lm\_no\_inter substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade + gender   
clm\_lm\_sig\_inter substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade + gender + gender \* adult\_ss  
 link: threshold:  
clm\_lm\_no\_inter logit flexible   
clm\_lm\_sig\_inter logit flexible   
  
 no.par AIC logLik LR.stat df Pr(>Chisq)  
clm\_lm\_no\_inter 8 18600 -9291.8   
clm\_lm\_sig\_inter 9 18602 -9291.8 0.0099 1 0.9207

## 5.3 Stratified analysis: gendered models

formula: substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade  
data:   
ordinalRegres\_substanceUse[ordinalRegres\_substanceUse$gender == "Male", ]  
  
 link threshold nobs logLik AIC niter max.grad cond.H   
 logit flexible 17969 -4857.70 9729.40 6(0) 3.65e-07 2.7e+04  
  
Coefficients:  
 Estimate Std. Error z value Pr(>|z|)   
adult\_ssYes -0.855355 0.107664 -7.945 1.95e-15 \*\*\*  
peer\_ssYes 0.265738 0.091771 2.896 0.003784 \*\*   
teacher\_ssYes -0.239755 0.062510 -3.835 0.000125 \*\*\*  
ptsd\_total 0.085324 0.003874 22.025 < 2e-16 \*\*\*  
grade 0.405034 0.019377 20.903 < 2e-16 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Threshold coefficients:  
 Estimate Std. Error z value  
None|Low 6.5656 0.2265 28.98  
Low|High 7.8100 0.2315 33.74

formula: substanceUse ~ adult\_ss + peer\_ss + teacher\_ss + ptsd\_total + grade  
data:   
ordinalRegres\_substanceUse[ordinalRegres\_substanceUse$gender == "Female", ]  
  
 link threshold nobs logLik AIC niter max.grad cond.H   
 logit flexible 18516 -4415.85 8845.71 7(0) 8.06e-13 4.1e+04  
  
Coefficients:  
 Estimate Std. Error z value Pr(>|z|)   
adult\_ssYes -0.894493 0.107156 -8.348 < 2e-16 \*\*\*  
peer\_ssYes 0.239199 0.106957 2.236 0.0253 \*   
teacher\_ssYes -0.292436 0.065182 -4.486 7.24e-06 \*\*\*  
ptsd\_total 0.072182 0.003688 19.572 < 2e-16 \*\*\*  
grade 0.367453 0.020466 17.955 < 2e-16 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Threshold coefficients:  
 Estimate Std. Error z value  
None|Low 6.3645 0.2494 25.52  
Low|High 7.9709 0.2566 31.07

# 6. Descriptive stats table

# Turn grade into a factor  
ordinalRegres\_substanceUse$grade <- factor(ordinalRegres\_substanceUse$grade,  
 labels = c(  
 "Grade 7",  
 "Grade 8",  
 "Grade 9",  
 "Grade 10",  
 "Grade 11",  
 "Grade 12"  
 )  
)  
  
# substanceUse: Check for NA values and convert them to an explicit factor level  
if (any(is.na(ordinalRegres\_substanceUse$substanceUse))) {  
 ordinalRegres\_substanceUse$substanceUse <- forcats::fct\_na\_value\_to\_level(ordinalRegres\_substanceUse$substanceUse)  
}  
  
# Create table  
reset\_gtsummary\_theme()  
theme\_gtsummary\_compact()  
  
table\_included <-  
 ordinalRegres\_substanceUse %>%  
 dplyr::select(  
 region,  
 grade,  
 gender,  
 ptsd\_total,  
 social\_support\_total,  
 adult\_ss,  
 peer\_ss,  
 teacher\_ss,  
 substanceUse  
 ) %>%  
 tbl\_summary(  
 by = substanceUse,  
 statistic = ptsd\_total ~ "{mean} ({sd})",  
 digits = list(  
 all\_continuous() ~ c(3, 2),  
 all\_dichotomous() ~ c(0, 2),  
 all\_categorical() ~ c(0, 2)  
 ),  
 missing = "always",  
 missing\_text = "Missing values",  
 type = list(  
 region ~ "categorical",  
 grade ~ "categorical",  
 gender ~ "categorical",  
 ptsd\_total ~ "continuous",  
 social\_support\_total ~ "categorical",  
 adult\_ss ~ "categorical",  
 peer\_ss ~ "categorical",  
 teacher\_ss ~ "categorical"  
 ),  
 label = list(  
 region = "Region",  
 grade = "Grade",  
 gender = "Gender",  
 ptsd\_total = "PTSD Total",  
 social\_support\_total = "Social Support Total",  
 adult\_ss = "Adult SS",  
 peer\_ss = "Peer SS",  
 teacher\_ss = "Teacher SS"  
 ),  
 # sort = region ~ "frequency",  
 percent = "column"  
 ) %>%  
 add\_n() %>%  
 bold\_labels() %>%  
 modify\_spanning\_header(all\_stat\_cols() ~ "\*\*Substance Use Level\*\*")  
  
table\_included %>%  
 as\_flex\_table() %>%  
 fontsize(size = 10, part = "all") %>%  
 save\_as\_docx(path = here("output", "table\_summary\_regresSample.docx"))  
  
table\_included

|  | | **Substance Use Level** | | |
| --- | --- | --- | --- | --- |
| **Characteristic** | **N** | **None**, N = 34,0751 | **Low**, N = 1,7411 | **High**, N = 6691 |
| **Region** | 36,485 |  |  |  |
| Arecibo |  | 1,906 (5.59%) | 71 (4.08%) | 21 (3.14%) |
| Bayamon |  | 4,598 (13.49%) | 214 (12.29%) | 60 (8.97%) |
| Caguas |  | 4,455 (13.07%) | 274 (15.74%) | 95 (14.20%) |
| Humacao |  | 5,285 (15.51%) | 282 (16.20%) | 133 (19.88%) |
| Mayaguez |  | 6,421 (18.84%) | 303 (17.40%) | 104 (15.55%) |
| Ponce |  | 6,811 (19.99%) | 348 (19.99%) | 131 (19.58%) |
| San Juan |  | 4,599 (13.50%) | 249 (14.30%) | 125 (18.68%) |
| Missing values |  | 0 | 0 | 0 |
| **Grade** | 36,485 |  |  |  |
| Grade 7 |  | 6,609 (19.40%) | 104 (5.97%) | 41 (6.13%) |
| Grade 8 |  | 6,422 (18.85%) | 146 (8.39%) | 52 (7.77%) |
| Grade 9 |  | 5,500 (16.14%) | 222 (12.75%) | 85 (12.71%) |
| Grade 10 |  | 5,476 (16.07%) | 304 (17.46%) | 106 (15.84%) |
| Grade 11 |  | 5,468 (16.05%) | 415 (23.84%) | 160 (23.92%) |
| Grade 12 |  | 4,600 (13.50%) | 550 (31.59%) | 225 (33.63%) |
| Missing values |  | 0 | 0 | 0 |
| **Gender** | 36,485 |  |  |  |
| Male |  | 16,694 (48.99%) | 860 (49.40%) | 415 (62.03%) |
| Female |  | 17,381 (51.01%) | 881 (50.60%) | 254 (37.97%) |
| Missing values |  | 0 | 0 | 0 |
| **PTSD Total** | 36,485 | 7.272 (6.48) | 11.573 (7.61) | 12.833 (9.53) |
| Missing values |  | 0 | 0 | 0 |
| **Social Support Total** | 36,485 |  |  |  |
| 0 of 3 Endorsed |  | 668 (1.96%) | 54 (3.10%) | 45 (6.73%) |
| 1 of 3 Endorsed |  | 2,930 (8.60%) | 189 (10.86%) | 82 (12.26%) |
| 2 of 3 Endorsed |  | 12,667 (37.17%) | 742 (42.62%) | 281 (42.00%) |
| 3 of 3 Endorsed |  | 17,810 (52.27%) | 756 (43.42%) | 261 (39.01%) |
| Missing values |  | 0 | 0 | 0 |
| **Adult SS** | 36,485 |  |  |  |
| No |  | 1,570 (4.61%) | 188 (10.80%) | 101 (15.10%) |
| Yes |  | 32,505 (95.39%) | 1,553 (89.20%) | 568 (84.90%) |
| Missing values |  | 0 | 0 | 0 |
| **Peer SS** | 36,485 |  |  |  |
| No |  | 4,252 (12.48%) | 205 (11.77%) | 113 (16.89%) |
| Yes |  | 29,823 (87.52%) | 1,536 (88.23%) | 556 (83.11%) |
| Missing values |  | 0 | 0 | 0 |
| **Teacher SS** | 36,485 |  |  |  |
| No |  | 14,709 (43.17%) | 889 (51.06%) | 366 (54.71%) |
| Yes |  | 19,366 (56.83%) | 852 (48.94%) | 303 (45.29%) |
| Missing values |  | 0 | 0 | 0 |
| 1n (%); Mean (SD) | | | | |