Fit Model

Ikramul H Khan

2025-07-19

The very first step we load libraries

```
# Load libraries
library(haven)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

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

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

```
library(survey)
```

```
## Loading required package: grid

## Loading required package: Matrix

## Loading required package: survival

##
## Attaching package: 'survey'

## The following object is masked from 'package:graphics':
##
##     dotchart
```

```
library(srvyr)
```

```
##
## Attaching package: 'srvyr'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(broom)
library(flextable)
library(officer)
```

Load dataset and modification for our analysis

```
# Open the dataset
BDHS <- read_sav("adolescent fertility new_1.SAV")
#Recode the adolescent fertility using new variable V213 as currenly pregnant status

BDHS <- BDHS %>%
  mutate(adol_fertility = ifelse(V201 >= 1 | V213 == 1, 1, 0))
BDHS
```

```
## # A tibble: 2,449 x 20
##    V013      V024       V025    V106    V130    V151    V701    WomenEmpowerment
##    <dbl+lbl> <dbl+lbl>  <dbl+l> <dbl+l> <dbl+l> <dbl+l> <dbl+l> <dbl+lbl>
##  1 1 [15-19] 1 [Barish~ 1 [Urb~ 2 [Sec~ 1 [Mus~ 1 [Mal~ 2 [Sec~ 1 [No]
##  2 1 [15-19] 1 [Barish~ 1 [Urb~ 2 [Sec~ 1 [Mus~ 1 [Mal~ 2 [Sec~ 1 [No]
##  3 1 [15-19] 1 [Barish~ 1 [Urb~ 1 [Pri~ 1 [Mus~ 2 [Fem~ 2 [Sec~ 1 [No]
##  4 1 [15-19] 1 [Barish~ 1 [Urb~ 2 [Sec~ 1 [Mus~ 1 [Mal~ 2 [Sec~ 1 [No]
##  5 1 [15-19] 1 [Barish~ 1 [Urb~ 1 [Pri~ 1 [Mus~ 1 [Mal~ 0 [No ~ 0 [Empowered]
##  6 1 [15-19] 1 [Barish~ 1 [Urb~ 0 [No ~ 1 [Mus~ 2 [Fem~ 1 [Pri~ 0 [Empowered]
##  7 1 [15-19] 1 [Barish~ 1 [Urb~ 2 [Sec~ 1 [Mus~ 1 [Mal~ 2 [Sec~ 1 [No]
##  8 1 [15-19] 1 [Barish~ 1 [Urb~ 2 [Sec~ 1 [Mus~ 1 [Mal~ 0 [No ~ 0 [Empowered]
##  9 1 [15-19] 1 [Barish~ 1 [Urb~ 2 [Sec~ 1 [Mus~ 1 [Mal~ 2 [Sec~ 1 [No]
## 10 1 [15-19] 1 [Barish~ 1 [Urb~ 2 [Sec~ 1 [Mus~ 1 [Mal~ 1 [Pri~ 1 [No]
## # i 2,439 more rows
## # i 12 more variables: V012 <dbl+lbl>, V190 <dbl+lbl>, V312New <dbl+lbl>,
## #   Age_Gap <dbl+lbl>, V201 <dbl>, CEB <dbl>, `filter_$` <dbl+lbl>, V001 <dbl>,
## #   V005 <dbl>, V023 <dbl+lbl>, V213 <dbl+lbl>, adol_fertility <dbl>
```

```
BDHS <- BDHS %>%
  mutate(
    adol_fertility = as.numeric(adol_fertility)
  )

# Prepare variables
BDHS <- BDHS %>%
  mutate(
    education = as_factor(V106),
    partner_education = as_factor(V701),
    division = as_factor(V024),
    residence = as_factor(V025),
    religion = as_factor(V130),
    wealth = as_factor(V190),
```

```
    age = as_factor(V012),
    age_gap = as_factor(Age_Gap),
    contraceptive_status = as_factor(V312New),
    WomenEmpowerment = as_factor(WomenEmpowerment),
    weight = V005 / 1000000
  )

# Relevel the reference category

BDHS <- BDHS %>%
  mutate(
    education = relevel(education, ref = "Higher"),
    partner_education = relevel(partner_education, ref = "Higher"),
    age_gap = relevel(age_gap, ref = "<=5"),
    wealth = relevel(wealth, ref = "Rich")
  )

options(survey.lonely.psu = "adjust")
# Create survey design object
bdhs_design <- BDHS %>%
  as_survey_design(
    ids = V001,
    strata = V023,
  weights = weight,
    nest = TRUE
  )
```

Check the overdispersion of the dataset

```
# 1. Fit survey-weighted logistic regression model and check overdispersion
model_binom <- svyglm(
  adol_fertility ~ education + partner_education + division + residence +
    religion + wealth + age + age_gap + contraceptive_status + WomenEmpowerment,
  design = bdhs_design,
  family = binomial()
)
```

```
## Warning in eval(family$initialize): non-integer #successes in a binomial glm!
```

```
# 2. Check model summary
summary(model_binom)
```

```
##
## Call:
## svyglm(formula = adol_fertility ~ education + partner_education +
##     division + residence + religion + wealth + age + age_gap +
##     contraceptive_status + WomenEmpowerment, design = bdhs_design,
##     family = binomial())
##
## Survey design:
```

```
## Called via srvyr
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   -2.99977    0.40637  -7.382 4.87e-08 ***
## educationNo education          1.01262    0.42834   2.364 0.025246 *
## educationPrimary               0.82911    0.25068   3.307 0.002591 **
## educationSecondary             0.54514    0.18263   2.985 0.005832 **
## partner_educationNo education  0.76663    0.30615   2.504 0.018376 *
## partner_educationPrimary       0.67315    0.17163   3.922 0.000518 ***
## partner_educationSecondary     0.32502    0.17374   1.871 0.071876 .
## divisionChattogram             0.48339    0.21898   2.207 0.035648 *
## divisionDhaka                  0.42309    0.21934   1.929 0.063928 .
## divisionKhulna                 0.59142    0.21641   2.733 0.010754 *
## divisionMymensingh             0.01308    0.24074   0.054 0.957041
## divisionRajshahi               0.24822    0.22214   1.117 0.273328
## divisionRangpur                0.40555    0.20744   1.955 0.060632 .
## divisionSylhet                 0.18970    0.19198   0.988 0.331556
## residenceRural                -0.03072    0.12591  -0.244 0.808998
## religionNon-Muslim            -0.04794    0.26817  -0.179 0.859397
## wealthPoor                     0.79125    0.15921   4.970 3.01e-05 ***
## wealthMiddle                   0.49786    0.15724   3.166 0.003709 **
## ageAge 16                      0.79832    0.21595   3.697 0.000941 ***
## ageAge 17                      1.29226    0.20790   6.216 1.03e-06 ***
## ageAge 18                      1.64608    0.21644   7.605 2.77e-08 ***
## ageAge 19                      2.05780    0.22259   9.245 5.29e-10 ***
## age_gap>=11                    0.79422    0.19695   4.033 0.000385 ***
## age_gap6-10                    0.17293    0.15282   1.132 0.267395
## contraceptive_statusNo        -0.05137    0.14887  -0.345 0.732614
## WomenEmpowermentNo            -0.08132    0.21068  -0.386 0.702402
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 0.9956332)
##
## Number of Fisher Scoring iterations: 4
```

```r
# 3. Check for overdispersion manually
res_dev <- deviance(model_binom)          # residual deviance
df_res <- df.residual(model_binom)        # degrees of freedom
dispersion <- res_dev / df_res            # dispersion statistic

# 4. Print dispersion value
dispersion
```

```
## [1] 108.3321
```

Fit the appropriate survey-weighted logistic model for over-disperse data

```r
# Fit weighted logistic for over-disperse data

model <- svyglm(
  adol_fertility ~ education + partner_education + division + residence +
    religion + wealth + age + age_gap + contraceptive_status + WomenEmpowerment,
  design = bdhs_design,
  family = quasibinomial()
)

# Tidy and exponentiate coefficients
results <- tidy(model, conf.int = TRUE, exponentiate = TRUE) %>%
  rename(
    AOR = estimate,
    `2.5%` = conf.low,
    `97.5%` = conf.high,
    `p-value` = p.value
  ) %>%
  select(term, AOR, `2.5%`, `97.5%`, `p-value`)


# Tidy and exponentiate coefficients
results <- results %>%
  rename(
    `Predictor` = term,
    `AOR` = AOR,
    `95% CI Lower` = `2.5%`,
    `95% CI Upper` = `97.5%`,
    `P-value` = `p-value`
  )

results <- results %>%
  mutate(across(where(is.numeric), ~ round(., 3)))
results
```

```
## # A tibble: 26 x 5
##    Predictor                       AOR `95% CI Lower` `95% CI Upper` `P-value`
##    <chr>                         <dbl>          <dbl>          <dbl>     <dbl>
##  1 (Intercept)                    0.05          0.022          0.114     0
##  2 educationNo education          2.75          1.14           6.62      0.025
##  3 educationPrimary               2.29          1.37           3.83      0.003
##  4 educationSecondary             1.72          1.19           2.51      0.006
##  5 partner_educationNo education  2.15          1.15           4.03      0.018
##  6 partner_educationPrimary       1.96          1.38           2.79      0.001
##  7 partner_educationSecondary     1.38          0.97           1.98      0.072
##  8 divisionChattogram             1.62          1.03           2.54      0.036
##  9 divisionDhaka                  1.53          0.974          2.39      0.064
## 10 divisionKhulna                 1.81          1.16           2.81      0.011
## # i 16 more rows
```

```r
# Create flextable
ft <- flextable(results) %>%
  autofit() %>%
  set_table_properties(width = 1, layout = "autofit") %>%
```

```
  theme_zebra() %>%
  align(align = "center", part = "all") %>%
  bold(i = ~ `P-value` < 0.05, bold = TRUE, part = "body")  # bold significant

# Create new Word doc and add table
doc <- read_docx() %>%
  body_add_par("Table X: Adjusted Odds Ratios for Adolescent Fertility", style = "heading 1") %>%
  body_add_flextable(ft)

# Save to Word file with new name
print(doc, target = "Adjusted_OR_Table_Adolescent_Fertility.docx")
```