

Model Analysis

Ikramul H Khan

2025-07-17

```
# 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
```

```
# Open the dataset
BDHS <- read_sav("adolescent fertility new_1.SAV")
#Recode the adolescent fertility using new variable V213 as currently 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+lbl> <dbl+lbl> <dbl+lbl> <dbl+lbl> <dbl+lbl> <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>
```

```
# 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
  )
```

```
# Create survey design object
```

```
bdhs_design <- BDHS %>%
  as_survey_design(
    ids = V001,
    strata = V023,
    weights = weight,
    nest = TRUE
  )
```

```
# For strata with single PSU
```

```
options(survey.lonely.psu = "adjust")
```

```
# Variables to summarize
```

```
vars_to_summarize <- c(
  "education", "partner_education", "division", "residence", "religion",
  "wealth", "contraceptive_status", "age", "age_gap", "WomenEmpowerment"
)

# Function to summarize each variable
get_summary <- function(var) {
  bdhs_design %>%
    group_by(value = .data[[var]]) %>%
    summarise(
      n = unweighted(n()),
      percent = survey_mean(proportion = TRUE, na.rm = TRUE) * 100,
      .groups = "drop"
    ) %>%
    mutate(variable = var) %>%
    select(variable, category = value, n, percent)
}

# Apply and combine
descriptive_table <- bind_rows(lapply(vars_to_summarize, get_summary))
```

```
## Warning: There was 1 warning in `dplyr::summarise()`.
## i In argument: `percent = survey_mean(proportion = TRUE, na.rm = TRUE) * 100`.
## i In group 1: `value = No education`.
## Caused by warning:
## ! na.rm argument has no effect on survey_mean when calculating grouped proportions.
## This warning is displayed once per session.
```

```
# View table
print(descriptive_table)
```

```
## # A tibble: 35 x 4
##   variable      category      n percent
##   <chr>         <fct>    <int>   <dbl>
## 1 education    No education    40    1.37
## 2 education    Primary       339   14.6
## 3 education    Secondary     1879   76.9
## 4 education    Higher        191    7.11
## 5 partner_education No education   117    4.85
## 6 partner_education Primary       429   17.9
## 7 partner_education Secondary     1541   62.9
## 8 partner_education Higher        362   14.3
## 9 division     Barishal      271    2.98
## 10 division     Chattogram    340    5.36
## # i 25 more rows
```

```
# Export as CSV
write.csv(descriptive_table, "descriptive_table.csv", row.names = FALSE)

# Load libraries
library(officer)
```

```

library(flextable)

# Create Word document and add formatted table
doc <- read_docx() %>%
  body_add_par("Descriptive Table: Weighted Percentage and Unweighted Frequency", style = "heading 1") %>%
  body_add_flextable(flextable(descriptive_table))

# Fix lonely PSU issue
options(survey.lonely.psu = "adjust")
library(tidyr)

##
## Attaching package: 'tidyr'

## The following objects are masked from 'package:Matrix':
##
##      expand, pack, unpack

# Background variables
vars_to_check <- c(
  "education", "partner_education", "division", "residence", "religion",
  "wealth", "contraceptive_status", "age", "age_gap", "WomenEmpowerment"
)

# Function to create table for each variable
generate_table <- function(varname) {
  # Unweighted counts
  unweighted_counts <- BDHS %>%
    filter(!is.na(.data[[varname]]), !is.na(adol_fertility)) %>%
    count(!sym(varname), adel_fertility) %>%
    pivot_wider(names_from = adel_fertility, values_from = n, values_fill = 0) %>%
    rename(`No Child (n)` = `0`, `Had Child (n)` = `1`) %>%
    mutate(Category = as.character(!sym(varname))) %>%
    select(Category, `Had Child (n)`, `No Child (n)`)

  # Weighted percentages
  tab <- svytable(as.formula(paste("~", varname, "+ adel_fertility")), bdhs_design)
  prop_tab <- prop.table(tab, margin = 1) * 100
  perc_tab <- round(as.data.frame.matrix(prop_tab), 1)
  perc_tab$Category <- rownames(perc_tab)
  rownames(perc_tab) <- NULL
  names(perc_tab) <- c("No Child (%)", "Had Child (%)", "Category")

  # Chi-square p-value
  pval <- tryCatch({
    svychisq(as.formula(paste("~", varname, "+ adel_fertility")), bdhs_design)$p.value
  }, error = function(e) NA)
  pval_str <- format.pval(pval, digits = 3)

  # Merge and format final table
  final <- left_join(unweighted_counts, perc_tab, by = "Category") %>%
    mutate(

```

```

    Variable = varname,
    `p-value` = c(pval_str, rep("", n() - 1))
  ) %>%
  select(Variable, Category,
    `Had Child (n)`, `Had Child (%)`,
    `No Child (n)`, `No Child (%)`,
    `p-value`)

  return(final)
}

# Apply function across all variables
full_table <- bind_rows(lapply(vars_to_check, generate_table))

# Export to Word document
doc <- read_docx() %>%
  body_add_par("Frequency (%) Distribution of Ever-Married Adolescent Women by Fertility with p-value (")
  body_add_flextable(flextable(full_table))

print(doc, target = "Fertility_Distribution_by_Covariates_new.docx")

```