BDA - Project

Anonymous

Contents


```
knitr::opts_chunk$set(echo = TRUE, fig.height = 6, fig.width = 10)
# load libraries
pkgs <- c("tidyr", "ggplot2", "rstan", "bayesplot", "parallel", "loo",
          "readxl")
sapply(pkgs, require, character.only = TRUE)
```

```
##     tidyr   ggplot2     rstan bayesplot  parallel       loo    readxl
##      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE      TRUE
```

```
# set number of default cores
options(mc.cores = detectCores())
# set plot theme
theme_set(theme_minimal())
# set colour scheme
color_scheme_set("blue")
```

Note: this is the contents page. The assignment starts on the following page.

1. Introduction

This project is motivated by a thesis I finished in 2020. The thesis included a pairwise meta-analysis to estimate the efficacy of the Bivalent Human Papillomavirus vaccine. It was originally coded in JAGS, did not have any proper priors, did not include coefficients, and it was used to inform a fully-integrated Bayesian health economics model. The original JAGS model was coded as:

```
# likelihood and prior. Model parameters are abbreviated by .vac.
  for (i in 1:Nstud.vac) {
    # Likelihood:
     rA.vac[i] ~ dbin(pA.vac[i], nA.vac[i])
     rB.vac[i] ~ dbin(pB.vac[i], nB.vac[i])

    # Logistic link function:
     logit(pA.vac[i]) <- mu.vac[i]
     logit(pB.vac[i]) <- mu.vac[i] + delta.vac[i]

    # Average effect prior for SUB-MODEL 2:
     mu.vac[i] ~ dnorm(0, 1e-4)
    # Prior for sub-model 2 (Random. pop. effect):
     delta.vac[i] ~ dt(psi.vac, prec.vac, 1)
     # if desired can be ~ dnorm(psi.vac, prec.vac)

     ### Mixed predictive check for SUB-MODEL 2:
       # Predictive likelihood:
        rA.mxd[i] ~ dbin(pA.new[i], nA.vac[i])

       # Predictive logit link function:
        logit(pA.new[i]) <- mu.vac[i] + delta.new

       # Mixed predictve p-value:
        pA.mxd[i] <- step(rA.mxd[i] - rA.vac[i]) - 0.5 * equals(rA.mxd[i], rA.vac[i])
```

Since finishing the thesis, I have wanted toredo the meta-analysis in Stan. Note that this project won't include the economic model. Please see a simple illustration of the model below (note that some of the final model assumptions may differ).

Given the above motivation, I will therefore implement the original model in Stan and include proper priors, which will be given logical boundaries when necessary. Coefficients will also be included where applicable. I will also compare grouped and hierarchical models.

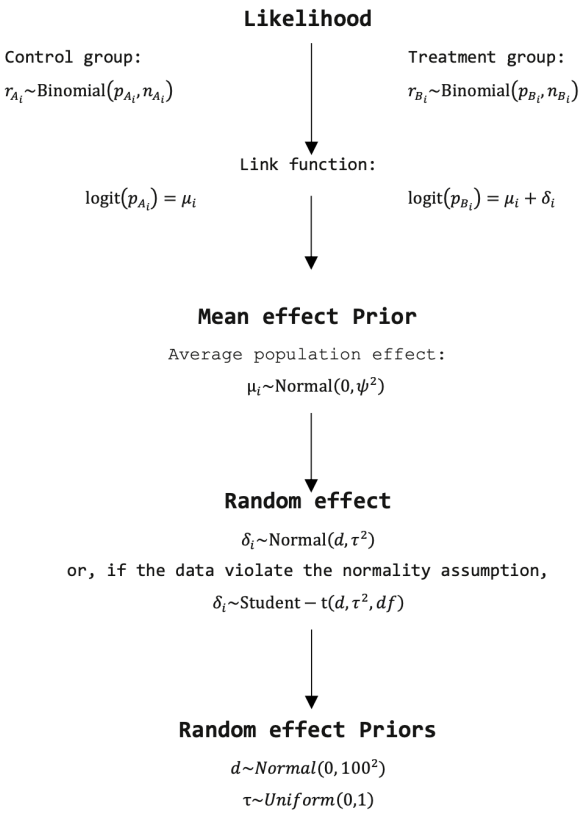

Figure 1: Model illustration

1.2 Data

The data for the model are printed below. Note that references for each input can be found in the original Excel file.

```
# load data
df <- read_excel("data/data_vaccine_case_control.xlsx")
```

```
## New names:
## * `` -> `...5`
```

```
# print data frame
df
```

```
## # A tibble: 11 x 5
##    y_vaccine n_vaccine y_control n_control ...5
##        <dbl>     <dbl>     <dbl>     <dbl> <chr>
##  1        32      5406       435      5375 Apter et al. (2015). Efficacy of Hum~
##  2        12       366        41       355 Harper et al. (2004). Efficacy of a ~
##  3         1       310        28       277 Harper et al. (2006). Sustained effi~
##  4         3      2190        38      2239 Herrero et al. (2011). Prevention of~
##  5        61      2910       219      2924 Herrero et al. (2013). Reduced Preva~
##  6         0       387        15       392 Konno et al. (2010). Efficacy of Hum~
##  7         0       193        10       175 Naud et al. (2014). Sustained effica~
##  8         2      7788        21      7838 Paavonen et al. (2007). Efficacy of ~
##  9         4      7344        56      7838 Paavonen et al. (2009). Efficacy of ~
## 10         0       401        20       372 Romanowski et al. (2009). Sustained ~
## 11         1      2497        17      2502 Zhu et al. (2014). Efficacy, immunog~
```