BDA - Project

Anonymous

Contents

```r
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
```

```r
# 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 random-effects meta-analysis to estimate the efficacy of the bivalent Cervarix Human Papillomavirus vaccine. It was originally coded in JAGS, did not have any proper priors, did not include coefficients, and was ultimately 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 random effect prior for SUB-MODEL 2:
     mu.vac[i] ~ dnorm(0, 1e-4)
    # Prior for sub-model 2 (ATE 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 to redo the meta-analysis in Stan. Note that this project won't, however, 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 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 a separate and hierarchical model.

Likelihood

Control group:

$$r_{A_i} \sim \text{Binomial}(p_{A_i}, n_{A_i})$$

Treatment group:

$$r_{B_i} \sim \text{Binomial}(p_{B_i}, n_{B_i})$$

Link function:

$$\text{logit}(p_{A_i}) = \mu_i$$

$$\text{logit}(p_{B_i}) = \mu_i + \delta_i$$

Mean effect Prior

Average population effect:

$$\mu_i \sim \text{Normal}(0, \psi^2)$$

Random effect

$$\delta_i \sim \text{Normal}(d, \tau^2)$$

or, if the data violate the normality assumption,

$$\delta_i \sim \text{Student} - \text{t}(d, \tau^2, df)$$

Random effect Priors

$$d \sim Normal(0, 100^2)$$
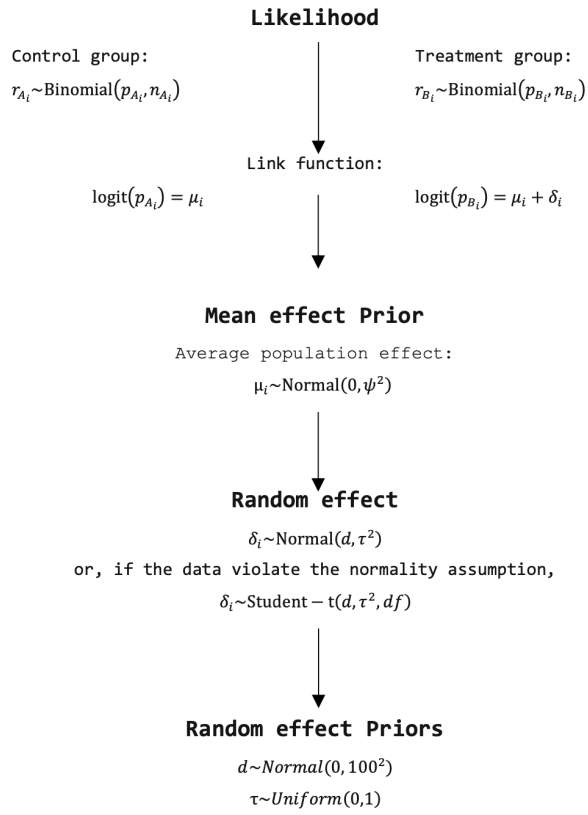
$$\tau \sim Uniform(0, 1)$$

Figure 1: Model illustration. Note that some parameter distributions for the final are different, since informative priors have been added

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
print(df)
```

```
## # A tibble: 11 x 5
##     y_vaccine n_vaccine y_control n_control ...5
##         <dbl>     <dbl>     <dbl>     <dbl> <chr>
## 1          22      5331       212      5291 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~
```

We can then prepare the data for Stan as follows:

```
# n studies
n_s <- nrow(df)
# y events in control arm
y_0 <- df$y_control
# y events in vaccine arm
y_1 <- df$y_vaccine
# n observations in control arm
n_0 <- df$n_control
# n observations in vaccine arm
n_1 <- df$n_vaccine
# data list
data_list <- list(n_s = n_s, y_0 = y_0, n_0 = n_0, y_1 = y_1, n_1 = n_1)
```

The model will thus loop through an $i \times j$ matrix, with i rows and j columns, where there are i studies with j arms.

3. Model

The general model is a simple binomial model with several priors and hyperiors (same as above). However, to provide an understanding of whether a separate or hierarchical model is better suited to the problem at

hand, we can ought to run the model as both a separate or hierarchical model, respectively. These different approaches to the model have a slightly different structure, which is detailed in the '*Model background and results*' section below.

3.1 Model background and results

Below is a brief discussion and illustration of each model.

3.3.1 The Separate model

The separate model is the simpler of the two types (in terms of the number of parameters). Notationally, the model can be represented as

$$y_{ij} \sim Bin(n_{ij}, p_{ij})$$

where y_{ij} and n_{ij} are the i'th events and observations for jth'th arm (vaccine or control), respectively. Then, to define a generalised linear model, we can use a logit link to develop a logistic regression model. Thus,

$$logit(p_{ij}) \equiv log(\frac{p_{ij}}{1 - pij})$$

which, more precisely, is the linear predictor

$$logit(p_{ij}) = \mu_i + \delta_{ij}$$

where

$$\mu_i \sim \mathcal{N}(-1.38, 1000)$$

and

$$\delta_{.j} = \begin{cases} 0, & \text{if } j = 0 \\ \sim Student - t(\nu, \psi, \tau), & \text{if } j = 1 \end{cases}$$

since the control arm is no average treatment effect (ATE). Note that μ_i represents the group-level 'random effect' and $j = 0$ and $j = 1$ represent the indexing for the control and vaccine arms respectively. The probability of efficacy is therefore assumed to be a additive combination of a random population effect and the average treatment effect, at a group-level, when applicable (conditional on whether the group arm received the treatment or not). Here, ψ represents a reasonable informative mean prior value. Note that the priors for μ_i and δ_{ij} in the separate model are recoded as hyperpriors in the hierarchical model.

Discussion on these values is therefore deferred to section 3.3.2. In addition, the within group treatment effect has been specified as a student-t distribution, to account for greater uncertainty in distribution tails (as there is large variation in the size of the studies and the group results). Finally, note that the Stan code for the separate model is coded as

```
// Random Effects
// Pair-wise meta-analysis
// Binomial model, logit link
data {
 int <lower = 0> n_s;
 int <lower = 0> y_0[n_s];
 int <lower = 0> n_0[n_s];
 int <lower = 0> y_1[n_s];
 int <lower = 0> n_1[n_s];
```

```
}
parameters {
 real mu[n_s];
 real delta[n_s];
}
transformed parameters {
 real p_eff[n_s];
 real odds_eff[n_s];
 for (i in 1:n_s) {
  // anti-logit
  odds_eff[i] = exp(delta[i]);      // transform log-odds ratio to odds-ratio
  p_eff[i] = 1 / (1 + odds_eff[i]); // transform odds ratio to probability
 }
}
model {
 // priors
 for (i in 1:n_s) {
  // group-level random effect (Matthijsse S. et al)
  mu[i] ~ normal(-1.38, 1000);
  // group-level average treatment effect (ATE) (Kamolratanakul S. et al)
  delta[i] ~ student_t(4, 1.1, 1000);
  // control: binomial likelihood with logit link
  y_0[i] ~ binomial_logit(n_0[i], mu[i]);
  // vaccine: binomial likelihood with logit link
  y_1[i] ~ binomial_logit(n_1[i], mu[i] + delta[i]);
 }
}
generated quantities {
 real log_lik[n_s];
 for (i in 1:n_s) {
  // log-lik for treatment arm groups
  log_lik[i] = binomial_logit_lpmf(y_1[i] | n_1[i], mu[i] + delta[i]);
 }
}
// End file
```

The results of the model, for the probability of vaccine effectiveness, for each group i are

```
# print results
print(project_sep, pars = c("p_eff[1]", "p_eff[2]", "p_eff[3]", "p_eff[4]",
                            "p_eff[5]", "p_eff[6]", "p_eff[7]", "p_eff[8]",
                            "p_eff[9]", "p_eff[10]", "p_eff[11]"))
```

```
## Inference for Stan model: project_sep.
## 4 chains, each with iter=2000; warmup=1000; thin=1;
## post-warmup draws per chain=1000, total post-warmup draws=4000.
##
##           mean se_mean   sd 2.5%  25%  50%  75% 97.5% n_eff Rhat
## p_eff[1]  0.91    0.00 0.02 0.88 0.90 0.91 0.92  0.94   185 1.01
## p_eff[2]  0.80    0.00 0.05 0.70 0.77 0.80 0.83  0.88   167 1.04
## p_eff[3]  0.97    0.00 0.03 0.89 0.96 0.98 0.99  1.00   293 1.01
## p_eff[4]  0.93    0.00 0.04 0.83 0.91 0.94 0.96  0.98   203 1.01
## p_eff[5]  0.79    0.00 0.02 0.75 0.78 0.79 0.81  0.84   167 1.06
```

```
## p_eff[6]   1.00    0.00 0.00 1.00 1.00 1.00 1.00  1.00  2002 1.00
## p_eff[7]   1.00    0.00 0.00 1.00 1.00 1.00 1.00  1.00  3406 1.00
## p_eff[8]   0.91    0.01 0.06 0.77 0.87 0.92 0.95  0.99   137 1.04
## p_eff[9]   0.93    0.00 0.03 0.86 0.91 0.94 0.96  0.98   189 1.02
## p_eff[10]  1.00    0.00 0.00 1.00 1.00 1.00 1.00  1.00  2019 1.00
## p_eff[11]  0.95    0.00 0.05 0.82 0.93 0.96 0.99  1.00   193 1.04
##
## Samples were drawn using NUTS(diag_e) at Wed May 18 21:04:42 2022.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

Note: to avoid the verbose messaging from Stan's parser, I have excluded the execution code for both the separate and hierarhical model.

3.3.2 The Hierarchical model

Unlike the separate model, the hierarchical model includes a population-level average treatment effect (ATE) parameter ψ. Hence, unlike the previous model, here ψ is drawn from a $\mathcal{N}(\mu, \sigma^2)$ with mean μ and variance σ^2. Thus, the symbolic illustration of the hierarchical model differs slightly. The likelihood and priors are the same, so

$$y_{ij} \sim Bin(n_{ij}, p_{ij})$$
$$logit(p_{ij}) = \mu_i + \delta_{ij}$$

where

$$\delta_{\cdot j} = \begin{cases} 0, & \text{if } j = 0 \\ \sim Student - t(\nu, \psi, \tau), & \text{if } j = 1 \end{cases}$$

However, hyperiors are included to capture the expected variation of the '*population-level*' vaccine efficacy. This thus shrinks the estimate to towards a common value and, theoretically, reduces uncertainty. These hyperpriors are represented as

$$\psi \sim \mathcal{N}(-1.38, 1000)$$

and

$$\tau \sim Inv - \chi_2(5)$$

The values for the population-level average treatment effect parameter ψ is a skeptical prior. The intention is to pull the mean of the posterior distribution towards a more conservative estimate. Hence, the hyperprior ψ takes on a mean probability of $\pi = 0.75$. This translates to an odds of $\frac{\pi}{1-\pi} = 3$ and a log-odds of $\ln 3 \approx 1.1$. The degrees of freedom value for the $Inv - \chi_2$ distribution is set at 5, giving the parameter a reasonable range to vary over for the standard deviation of the population level distribution.

The group-level mean random effect μ_i is set to an informative prior and is intended to represent the probability of natural evasion experienced by women (see reference: Matthijsse S. et al). Hence, rather than being set to an average effect of having 0 immunity at baseline, the point estimate for the mean of this distribution is to a probability of $\zeta = 0.2$. Again, we have to first translate this value to odds $frac\zeta 1 - \zeta = 0.25$ and then to a log-odds scale, so that $\ln 0.25 \approx -1.38$.

Parameters ψ and μ_i also have respective standard deviation parameters each set to a point that allows these parameter to vary over a reasonable range. This reasoning was based on several iterations of both the

separate and hierarhical model, showing the need for greater dispersion. I assume that this is due to the variation in the sizes of the number of observations between the 11 studies used for this analysis. Lastly, you can find the hierarchical model's Stan syntax and results below.

```
// Random Effects
// Pair-wise meta-analysis
// Binomial model, logit link
data {
 int <lower = 0> n_s;
 int <lower = 0> y_0[n_s];
 int <lower = 0> n_0[n_s];
 int <lower = 0> y_1[n_s];
 int <lower = 0> n_1[n_s];
}
parameters {
 real psi;
 real mu[n_s];
 real delta[n_s];
 real <lower = 0, upper = 20> tau;
}
transformed parameters {
 real p_eff;
 real odds_eff;
 // anti-logit
 odds_eff = exp(psi);         // transform log-odds ratio to odds-ratio
 p_eff = 1 / (1 + odds_eff); // transform odds ratio to probability
}
model {
 // hyperpriors
 psi ~ normal(1.1, 1000);   // pop-level (log-odds) ATE (Kamolratanakul S. et al)
 tau ~ inv_chi_square(5);   // populatiuon-level stdev
 // priors
 for (i in 1:n_s) {
  // group-level random (log-odds) effect (Matthijsse S. et al)
  mu[i] ~ normal(-1.38, 1000);
  // group-level average treatment effect (ATE)
  delta[i] ~ student_t(4, psi, tau);
  // control: binomial likelihood with logit link
  y_0[i] ~ binomial_logit(n_0[i], mu[i]);
  // vaccine: binomial likelihood with logit link
  y_1[i] ~ binomial_logit(n_1[i], mu[i] + delta[i]);
 }
}
generated quantities {
 real log_lik[n_s];
 for (i in 1:n_s) {
  // log-lik for treatment arm groups
  log_lik[i] = binomial_logit_lpmf(y_1[i] | n_1[i], mu[i] + delta[i]);
 }
}
// End file
```

```
# print results
print(project_hrchl, pars = c("p_eff", "odds_eff", "psi", "tau"))
```

```
## Inference for Stan model: project_hrchl.
## 4 chains, each with iter=2000; warmup=1000; thin=1;
## post-warmup draws per chain=1000, total post-warmup draws=4000.
##
##           mean se_mean   sd  2.5%   25%   50%   75% 97.5% n_eff Rhat
## p_eff     0.92    0.00 0.02  0.87  0.90  0.92  0.93  0.96  2300 1.00
## odds_eff  0.09    0.00 0.03  0.04  0.07  0.09  0.11  0.14  2345 1.00
## psi      -2.47    0.01 0.32 -3.17 -2.65 -2.44 -2.25 -1.93  2082 1.00
## tau       0.51    0.01 0.25  0.19  0.33  0.46  0.63  1.12   808 1.01
##
## Samples were drawn using NUTS(diag_e) at Wed May 18 21:05:09 2022.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

4 Inspecting the models

The following section investigates the internal validity of both the separate model by implementing PSIS leave-one-out cross-validation using the `loo` package.

4.1 Separate model

```
# print results
loo(project_sep)
```

```
##
## Computed from 4000 by 11 log-likelihood matrix
##
##          Estimate  SE
## elpd_loo    -22.1 4.7
## p_loo         5.3 1.1
## looic        44.2 9.4
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                         Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)      0    0.0%   <NA>
##  (0.5, 0.7]   (ok)        1    9.1%   24
##    (0.7, 1]   (bad)       7   63.6%   9
##    (1, Inf)   (very bad)  3   27.3%   2002
## See help('pareto-k-diagnostic') for details.
```

```
plot(loo(project_hrchl), diagnostic = c("k"), label_points = TRUE,
     main = "PSIS diagnostic plot")
```

PSIS diagnostic plot
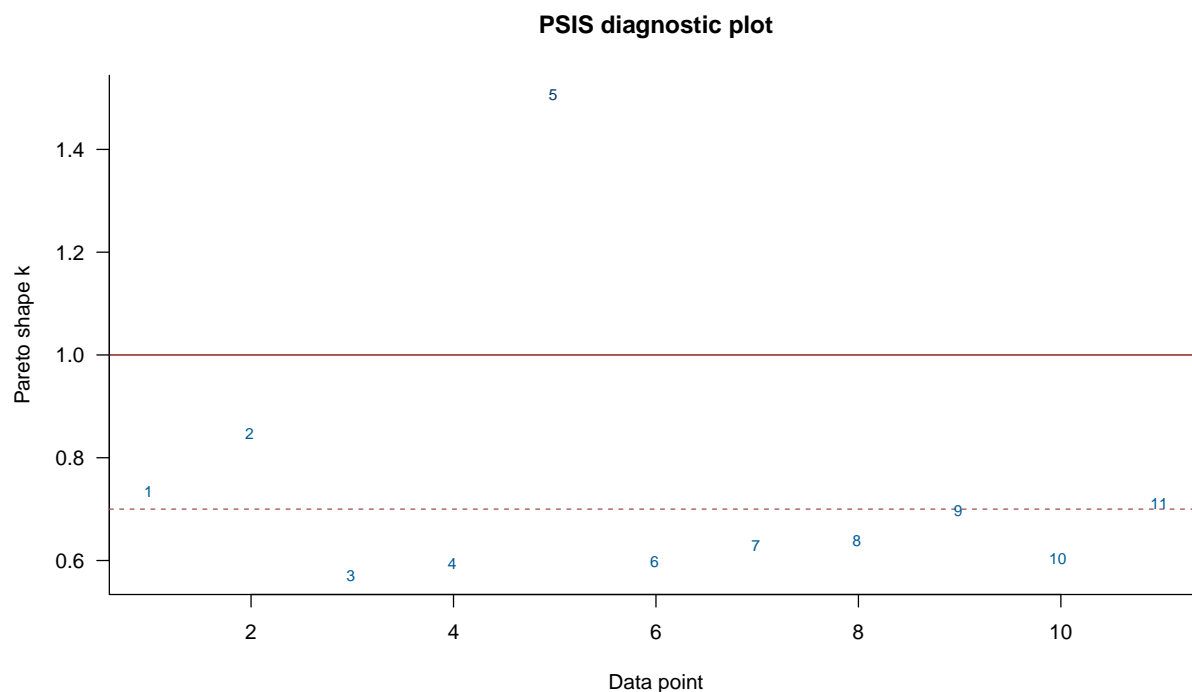
The above diagnostics indicate that the separate model failed to capture information for several of the groups, since many of the estimates are $\tilde{k} > 0.7$. In other words, given the $\tilde{k}$ results, this indicates that for some observations the leave-one-out posteriors are different enough from the full posterior that importance-sampling is not able to correct the difference. The model is not completely inadequately specified, given the `p_loo` result of 6.5 parameters. However, this is still roughly $2.5\times$ the number of parameters used in the model and so we can deduce that the model can serve as a building block for a more complex model.
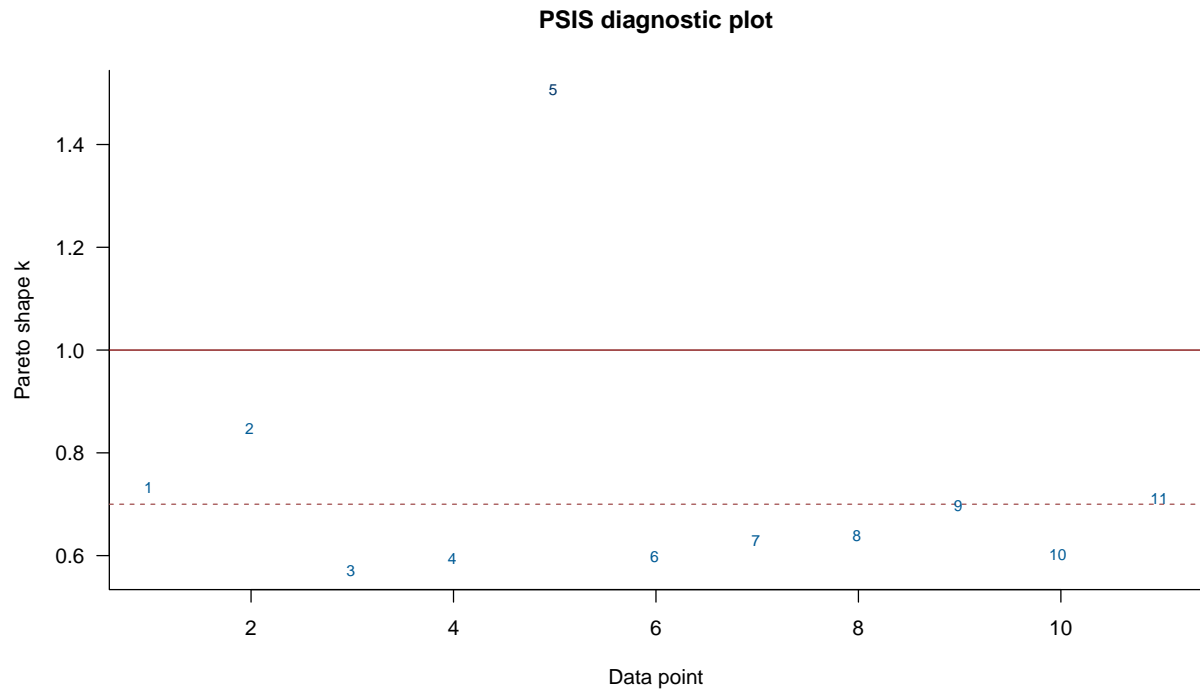
4.2 Hierarchical model

```
# print results
loo(project_hrchl)
```

```
##
## Computed from 4000 by 11 log-likelihood matrix
##
##          Estimate  SE
## elpd_loo    -26.3 5.0
## p_loo         7.2 2.4
## looic        52.5 9.9
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     0      0.0%   <NA>
##  (0.5, 0.7]   (ok)       7     63.6%   316
##    (0.7, 1]   (bad)      3     27.3%   71
##    (1, Inf)   (very bad) 1      9.1%   2
```

```
## See help('pareto-k-diagnostic') for details.
```

```
plot(loo(project_hrchl), diagnostic = c("k"), label_points = TRUE,
     main = "PSIS diagnostic plot")
```

PSIS diagnostic plot


```
loo_compare(list(
 "Separate Model" = loo(project_hrchl),
 "Hierarchical Model" = loo(project_sep))
 )
```

```
##                     elpd_diff se_diff
## Hierarchical Model  0.0       0.0
## Separate Model     -4.2       3.6
```

For the hierarchical model, the above indicate that the hierarchical is able to capture more information than the previous model. However, many of the estimates are still $\tilde{k} > 0.7$. Again, this indicates that for some observations the leave-one-out posteriors are different enough from the full posterior that importance-sampling is not able to correct the difference. The model is however, more adequately specified, given the `p_loo` result of only 6.5 parameters since 6 parameters we used to fit the model.

5. Discussion

From the above results and PSIS-LOO diagnostics, the hierarchical model has a more favourable expected log predictive density (ELPD) compared to the separate model. This implies that the hierarchical model can be expected to have a better predictive performance. In addition, the model indicates that the expected vaccine efficacy for the bivalent 16/18 Human Papillomavirus (HPV) is $\approx 92\%$. However, both models suffer from poor Pareto-$\tilde{k}$ values. Therefore, developing the hierarchical model further would be favourable. Incorporating coefficients, such as each study's latitude and or longitude may be beneficial. Another interesting estimate would be the value of information (VoI) of conducting an additional Randomised Control Trial (RCT) to further assess vaccine efficacy. Over and above these results, there is a need to consider 'double-counting' of some of the studies used in this analysis, as 3 of the RCTs used were long-term follow-ups of original papers.

6. Conclusion

A hierarchical model was used to synthesise 11 Randomised Control Trials to estimate the efficacy of the bivalent Cervarix HPV 16/18 strain vaccine. The expected posterior efficacy of the vaccine is estimated to be $\approx 92\%$, indicating that the bivalent 16/18 strain vaccine has a high-level of protection against the virus. However, the model suffers from poor Pareto-$\tilde{k}$ diagnostic values indicating that there is variation in the observed data that the model is unable to capture. Incorporating possible predictors, by fitting additional coefficients, would be beneficial. However, due to time constraints, this was not investigated further. Please see the Github repository to access the model code.