

STAT 447 Assignment 6

Caden Hewlett

2024-02-17

Question 1: Efficacy of Vaccines

In this exercise we will model the effectiveness of COVID vaccines using clinical trials data. Each trial consists of two arms: vaccinated (i.e., treated) and control (i.e., not treated). For a typical trial we know

- The total number of patients in each arm: t_v (vaccinated), t_c (control).
- The number of patients that got infected with the SARS-CoV-2 virus in each arm: n_v and n_c .

We model n_v and n_c as Binomial random variables. The unknown parameter for these distributions will depend on two numbers in $[0, 1]$

The first is Incidence (denoted p): the probability that a patient in the trial will become infected without being treated with the vaccine.

The second is Effectiveness (denoted e) the decrease in incidence that the vaccine provides.

We will take $\text{betaMP}(\mu, \lambda)$ distributions rather than $\text{beta}(\alpha, \beta)$. In this instance $\mu \in [0, 1]$ is the *mean* and $\lambda > 0$ is a *precision* parameter.

The following bijection holds regarding these parameters:

$$\mu = \frac{\alpha}{\alpha + \beta}, \text{ and } \lambda = \alpha + \beta \iff \alpha = \mu\lambda, \text{ and } \beta = (1 - \mu)\lambda$$

There are a few priors in this experiment.

Firstly, we have the Mean and Precision Priors of Effectiveness, given by a Uniform and Exponential distribution, respectively:

$$\mu_e \sim \text{unif}(0, 1), \text{ and } \lambda_e \sim \exp(0.001)$$

This gives us the Effectiveness, as a Likelihood given these Beta Parameters.

$$(e \mid \{\mu_e, \lambda_e\}) \sim \text{betaMP}(\mu_e, \lambda_e)$$

We then arrive at the number of infected inoculated individuals n_v , which arises as a function of the total number of individuals (the known value t_v), the incidence p (whose formulation will be discussed next) and e (discussed above).

$$(n_v \mid \{e, p\}) \sim \text{binom}(t_v, p(1 - e))$$

Now, we discuss the Bayesian framework for *all* individuals overall (i.e. “Incidence.”)

We have mean incidence and precision assigned priors as follows:

$$\mu_p \sim \text{betaMP}(0.1, 10), \text{ and } \lambda_p \sim \exp(0.001)$$

Similarly, we have the likelihood for incidence given these parameters:

$$(p \mid \{\mu_p, \lambda_p\}) \sim \text{betaMP}(\mu_p, \lambda_p)$$

Then, the number of infections in the control group:

$$(n_c \mid p) \sim \text{binom}(t_c, p)$$

Now, with this very robust framework presented, we will inspect the data before introducing a Hierarchical Model to describe the setting.

```
df = present = read.csv("vaccines.csv")
colnames(present) = c("Trials", "Arms", "Group Sizes", "Number of Cases")
kable(present)
```

Trials	Arms	Group Sizes	Number of Cases
Pfizer-BioNTech	vaccinated	18198	8
Pfizer-BioNTech	control	18325	162
Moderna-NIH	vaccinated	14134	11
Moderna-NIH	control	14073	185