MAYO CLINIC

Implementation Science and the Opportunity of Probabilistic Programming Languages

Patrick Wilson MPH[1], Curtis Storlie PhD[1]
1 Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery

Bayesian methods applied to implementation science can bridge the GAP between research and clinical practice

Background

Clinical research has failed to routinely translate evidence based practices into clinical care

The field of Dissemination and Implementation science (D&I) has emerged to study methods that increase the systematic uptake of research findings

A shift toward practice brings challenges:

Trials need to focus on effectiveness and be flexible enough to be done in clinic settings

Real world interventions tend to be imbedded in complex healthcare models with several components targeting different levels of the organization

Models should be sufficiently structured to capture variation to insure that relevant contextual factors are accounted for

Should be flexible enough to produce estimates that are more intuitive and meaningful to practice

Methods

To demonstrate the utility of Bayesian methods we present three cases studies

First, we show the estimation of a pragmatic clinical trial: The Stepped Wedge Design

Second, we estimate a mediation model

Third, we estimate measures on the additive scale and relative risk using logistic regression

Code is written in JAGS, Stan and Brms

Results

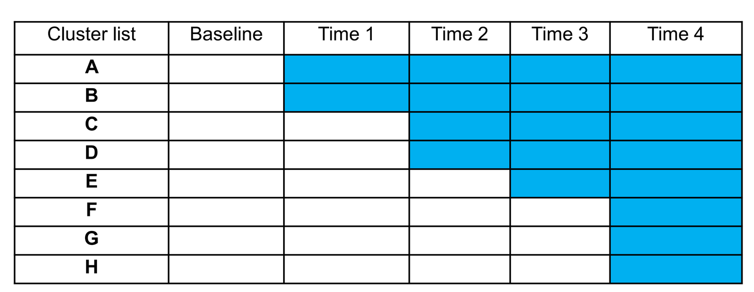
Pragmatic Clinical Trials

The Stepped Wedge Design is a cluster randomized trial where clusters crossover to the intervention in a random order

All clusters eventually receive the treatment

Mimics rollout structure which is common in clinical practice

$$Y_{ijk} = \mu + \beta_j + \alpha_i + \theta X_{ij} + e_{ijk}$$



The Stepped Wedge Design Schematic

```
```{r}
################
## JAGS Model ##
################
poisson_wedge <- "
model {
for(i in 1:N){
y[i] ~ dpois(lambda[i])
log(lambda[i]) <- mu[i]
mu[i] <- beta*treatment[i] + theta.clust[CID[i]] + u[TID[i]] + 1 * logoffset[i]
}

# random effects distributions (note: non-centered)
for (j in 1:Nprim) {
theta.clust[j] ~ dnorm(alpha, tau.clust)
RE.clust[j] <- theta.clust[j]-alpha
}

u[1] ~ dnorm(0, tau.time)
for (t in 2:Ntime){ u[t] ~ dnorm(rho * u[t - 1], tau.time)}

# priors on regression coefficients and variances
tau.clust ~ dgamma(1, 1) # between cluster variance
tau.time ~ dgamma(1, 0.05)  # time series variance

sigma2.clust <- 1/tau.clust
sigma2.time <- 1/tau.time

rho ~ dunif(0,1)  #Some AR(1) parameter

alpha ~ dnorm (0, 0.0001) # Intercept
beta ~ dnorm (0, 0.0001 ) # Treatment
}"
```
```

Clustering

Time effect

Mediation

Consists of two regression models. A mediator M model and outcome Y model for an intervention a adjusted by a covariate or set of covariates c.

$$E(M|a,c) = \beta_0 + \beta_1 a + \beta_2 c$$

$$E(Y|m,a,c) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 c$$

Where θ_1 is the "direct" effect and $\beta_1 * \theta_2$ is the "indirect" effect

Derivative Measures

Total effect

$$TE = DE + IE$$

Proportion mediated

$$PM = \frac{IE}{TE}$$

```
## Mediation Example
```{r, cache=TRUE}
#Original code from https://m-clark.github.io/posts/2019-03-12-mediation-models/
data(jobs)

model_mediator <- bf(job_seek ~ treat + econ_hard + sex + age)
model_outcome  <- bf(depress2 ~ treat + job_seek + econ_hard + sex + age)

med_result = brm(model_mediator + model_outcome + set_rescor(FALSE),
                 data = jobs,
                 cores=1)

summary(med_result)
print(sjstats::mediation(med_result), digits=4)

#Make stancode
#mediation
make_stancode(model_mediator + model_outcome + set_rescor(FALSE),
                 data = jobs)

#What changes if set_rescor(TRUE)
make_stancode(model_mediator + model_outcome + set_rescor(TRUE),
                                data = jobs)
```
```

Interaction and Post Processing

Traditionally many models tend to be on the relative scale (Cox, Poisson, logistic)

Several good reasons for this but often there is a need to have estimands that are clinically relevent

For example in epidemiology it is preferred to have tests for interaction on the additive scale vs. the relative scale

Interaction on the additive scale is straightforward

$$p_{11} - p_{10} - p_{01} + p_{00}$$

If $p11 - p10 - p01 + p00 > 0$ the interaction is said to be positive or "super-additive" If $p11 - p10 - p01 + p00 < 0$ the interaction is said to be negative or "sub-additive"

In a Bayesian framework this is a trival process which doesn't require switching models.

One just needs to transform your draws from the posterior, in this case from the log-odds scale to probability scale

The tricky part is the specfication of your population

```
```{r, cache=TRUE}

logistic_common <- brm(outcome_common ~ exposure*covariate,
                 data = risk_data,
                 family = bernoulli(),
                 iter = 2000)


make_stancode(outcome_common ~ exposure*covariate, data = risk_data)

```
```

```
#estimates of interest
risk_diff <- pr_1 - pr_0
RR <- pr_1 / pr_0

odds_1 <- pr_1/(1-pr_1)
odds_0 <- pr_0/(1-pr_0)

OR <- odds_1/odds_0

#test of interaction on the additive scale
interaction <- pr_11 - pr_10 - pr_01 + pr_00
```

Conclusions

Probabilistic programming languages like Stan and Jags make Bayesian modeling a viable alternative to *"off the shelf software"*

The ability to create flexible, module models can allow the elucidation of the complexities of practice

Although there is a learning curve with Bayesian models the knowledge is cumulative and will lead to more practice focused research

Key References

Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. *BMC Psychol.* 2015;3(1):32. Published 2015 Sep 16. doi:10.1186/s40359-015-0089-9

Stan Development Team. 2018. *Stan Modeling Language Users Guide and Reference Manual*, Version 2.18.0. http://mc-stan.org

Martyn Plummer (2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling, Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20–22, Vienna, Austria. ISSN 1609-395X.

Take a picture to access more content