Tests for Aphasia

Bayesian Approaches to Modeling the Conditional Dependence Between Multiple Diagnostic Tests

20 August, 2020

(Dendukuri & Joseph, 2001)

Dendukuri, N., & Joseph, L. (2001). Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. Biometrics, 57(1), 158-167.

N is the number of tests with subscript 11,10,01,00 representing $(T_1 = 1, T_2 = 1)$, $(T_1 = 1, T_2 = 0)$, $(T_1 = 0, T_2 = 1)$, $(T_1 = 0, T_2 = 0)$

Y is the true (latent) number of diseased

		T1		Total
		+	-	
T2	+	$N_{11}(Y_{11})$	$N_{01}(Y_{01})$	$N_{.1}(Y_{.1})$
	-	$N_{10}(Y_{10})$	$N_{00}(Y_{00})$	$N_{.0}(Y_{.0})$
Total		$N_{1.}(Y_{1.})$	$N_{0.}(Y_{0.})$	$N_{..}(Y_{..})$

		T1		Total
		+	-	
T2	+	2174(2125)	5076(4675)	7250(6800)
	-	26(0)	2724(1700)	2750(1700)
Total		2200(2125)	7800(6375)	10000(8500)

prevalence:$\hat{\pi} = \frac{Y_{..}}{N_{..}} = 0.85$;

sensitivity: $\hat{S}_1 = \frac{Y_{11}+Y_{10}}{Y_{..}} = 0.25$, $\hat{S}_2 = \frac{Y_{11}+Y_{01}}{Y_{..}} = 0.8$;

specificity: $\hat{C}_1 = \frac{N_{0.}-Y_{0.}}{N_{..}-Y_{..}} = 0.95$, $\hat{C}_2 = \frac{N_{1.}-Y_{1.}}{N_{..}-Y_{..}} = 0.7$

$u_s = \min(S_1, S_2) - S_1S_2 = S_1(1 - S_2) = 0.25 * (1 - 0.8) = 0.05$

$u_c = \min(C_1, C_2) - C_1C_2 = (1 - C_1)C_2 = (1 - 0.95) * 0.7 = 0.035$

$CovS \in [0, u_s] = [0, 0.05]$

$CovC \in [0, u_c] = [0, 0.035]$

$v_s = \max(S_1, S_2) - S_1S_2 = (1 - S_1)S_2 = (1 - 0.25) * 0.8 = 0.6$

$v_c = \max(C_1, C_2) - C_1C_2 = C_1(1 - C_2) = 0.95 * (1 - 0.7) = 0.285$

$S_1S_2 = 0.25 * 0.8 = 0.2$; $(1 - S_1)(1 - S_2) = 0.75 * 0.2 = 0.15$

$C_1C_2 = 0.95 * 0.7 = 0.665$; $(1 - C_1)(1 - C_2) = 0.05 * 0.3 = 0.015$

Fixed Effects Model

$$\begin{aligned}
&P(T_1=1,T_2=1|D=1)=S_1S_2+CovS_{12} && \in[0.2,0.2+0.05]=[0.2,0.25]\\
&P(T_1=1,T_2=0|D=1)=S_1(1-S_2)-CovS_{12} && \in[0.05-0.05,0.05]=[0,0.05]\\
&P(T_1=0,T_2=1|D=1)=(1-S_1)S_2-CovS_{12} && \in[0.6-0.05,0.6]=[0.55,0.6]\\
&P(T_1=0,T_2=0|D=1)=(1-S_1)(1-S_2)+CovS_{12} && \in[0.15,0.15+0.05]=[0.15,0.2]\\
&P(T_1=1,T_2=1|D=0)=(1-C_1)(1-C_2)+CovC_{12} && \in[0.015,0.015+0.035]=[0.015,0.05]\\
&P(T_1=1,T_2=0|D=0)=(1-C_1)C_2-CovC_{12} && \in[0.035-0.035,0.035]=[0,0.035]\\
&P(T_1=0,T_2=1|D=0)=C_1(1-C_2)-CovC_{12} && \in[0.285-0.035,0.285]=[0.25,0.285]\\
&P(T_1=0,T_2=0|D=0)=C_1C_2+CovC_{12} && \in[0.665,0.665+0.035]=[0.665,0.7]
\end{aligned}$$

the likelihood function of the observed data given the latent data

$$\begin{aligned}
L=&P(N_{11,10,01,00}|\pi,S_{1,2},C_{1,2},CovS_{12},CovC_{12},Y_{11,10,01,00})\\
&\propto\{\pi[S_1S_2+CovS_{12}]\}^{Y_{11}}\\
&\times\{\pi[S_1(1-S_2)-CovS_{12}]\}^{Y_{10}}\\
&\times\{\pi[(1-S_1)S_2-CovS_{12}]\}^{Y_{01}}\\
&\times\{\pi[(1-S_1)(1-S_2)+CovS_{12}]\}^{Y_{00}}\\
&\times\{[1-\pi][(1-C_1)(1-C_2)+CovC_{12}]\}^{N_{11}-Y_{11}}\\
&\times\{[1-\pi][(1-C_1)C_2-CovC_{12}]\}^{N_{10}-Y_{10}}\\
&\times\{[1-\pi][C_1(1-C_2)-CovC_{12}]\}^{N_{01}-Y_{01}}\\
&\times\{[1-\pi][C_1C_2+CovC_{12}]\}^{N_{00}-Y_{00}}
\end{aligned}$$

$\pi\sim Beta(\alpha_\pi,\beta_\pi)$

$S_j\sim Beta(\alpha_{S_j},\beta_{S_j})$, j=1,2

$C_j\sim Beta(\alpha_{C_j},\beta_{C_j})$, j=1,2

$0\le CovS_{12}\le\min(S_1,S_2)-S_1S_2=u_s$; $CovS_{12}\sim GBeta(\alpha_{covs12},\beta_{covs12})$

$0\le CovC_{12}\le\min(C_1,C_2)-C_1C_2=u_c$; $CovC_{12}\sim GBeta(\alpha_{covc12},\beta_{covc12})$

the joint posterior distribution of the parameters

$$\begin{aligned}
&P(\pi,S_{1,2},C_{1,2},CovS_{12},CovC_{12},Y_{11,10,01,00}|N_{11,10,01,00})\\
\propto\ &L\times\pi^{\alpha_\pi-1}(1-\pi)^{\beta_\pi-1}\\
&\times S_1^{\alpha_{S_1}-1}(1-S_1)^{\beta_{S_1}-1}S_2^{\alpha_{S_2}-1}(1-S_2)^{\beta_{S_2}-1}\\
&\times C_1^{\alpha_{C_1}-1}(1-C_1)^{\beta_{C_1}-1}C_2^{\alpha_{C_2}-1}(1-C_2)^{\beta_{C_2}-1}\\
&\times CovS_{12}^{\alpha_{covs12}-1}(u_s-CovS_{12})^{\beta_{covs12}-1}CovC_{12}^{\alpha_{covc12}-1}(u_c-CovC_{12})^{\beta_{covc12}-1}
\end{aligned}$$

- Full Conditional Distributions of Parameters in the Fixed Effects Model

$$\pi|N,Y_{11,10,01,00}\sim Beta\left(\alpha_\pi+\sum Y_{11,10,01,00},\beta_\pi+N-\sum Y_{11,10,01,00}\right)$$

$$
\begin{aligned}
&p(S_j|S_{3-j}, CovS_{12}, Y_{11,10,01,00}, \alpha_{S_j}, \beta_{S_j}, u_s, \beta_{covs12})\\
&\propto[S_1S_2 + CovS_{12}]^{Y_{11}}[S_1(1-S_2) - CovS_{12}]^{Y_{10}}[(1-S_1)S_2 - CovS_{12}]^{Y_{01}}[(1-S_1)(1-S_2) + CovS_{12}]^{Y_{00}}\\
&\quad\times S_j^{\alpha_{S_j}-1}(1-S_j)^{\beta_{S_j}-1}(u_s - CovS_{12})^{\beta_{covs12}-1}, j = 1,2\\
&p(C_j|C_{3-j}, CovC_{12}, N_{11,10,01,00}, Y_{11,10,01,00}, \alpha_{C_j}, \beta_{C_j}, u_c, \beta_{covc12})\\
&\propto[(1-C_1)(1-C_2) + CovC_{12}]^{N_{11}-Y_{11}}[(1-C_1)C_2 - CovC_{12}]^{N_{10}-Y_{10}}[C_1(1-C_2) - CovC_{12}]^{N_{01}-Y_{01}}[C_1C_2 + CovC_{12}]^{N_{00}-Y_{00}}\\
&\quad\times C_j^{\alpha_{C_j}-1}(1-C_j)^{\beta_{C_j}-1}(u_c - CovC_{12})^{\beta_{covc12}-1}, j = 1,2
\end{aligned}
$$

$$
\begin{aligned}
&p(CovS_{12}|S_{1,2}, Y_{11,10,01,00}, u_s, \alpha_{covs12}, \beta_{covs12})\\
&\propto[S_1S_2 + CovS_{12}]^{Y_{11}}[S_1(1-S_2) - CovS_{12}]^{Y_{10}}[(1-S_1)S_2 - CovS_{12}]^{Y_{01}}[(1-S_1)(1-S_2) + CovS_{12}]^{Y_{00}}\\
&\quad\times CovS_{12}^{\alpha_{covs12}-1}(u_s - CovS_{12})^{\beta_{covs12}-1}\\
&p(CovC_{12}|C_{1,2},, N_{11,10,01,00}, Y_{11,10,01,00}, u_c, \alpha_{covc12}, \beta_{covc12})\\
&\propto[(1-C_1)(1-C_2) + CovC_{12}]^{N_{11}-Y_{11}}[(1-C_1)C_2 - CovC_{12}]^{N_{10}-Y_{10}}[C_1(1-C_2) - CovC_{12}]^{N_{01}-Y_{01}}[C_1C_2 + CovC_{12}]^{N_{00}-Y_{00}}\\
&\quad\times CovC_{12}^{\alpha_{covc12}-1}(u_c - CovC_{12})^{\beta_{covc12}-1}
\end{aligned}
$$

$$
\begin{aligned}
&Y_{11}|\pi, S_{1,2}, C_{1,2}, CovS_{12}, CovC_{12}, N_{11} \sim Bino(N_{11}, p_{11})\\
\text{where } p_{11} = &\frac{\pi(S_1S_2 + CovS_{12})}{\pi(S_1S_2 + CovS_{12}) + (1-\pi)[(1-C_1)(1-C_2) + CovC_{12}]}\\
&Y_{10}|\pi, S_{1,2}, C_{1,2}, CovS_{12}, CovC_{12}, N_{10} \sim Bino(N_{10}, p_{10})\\
\text{where } p_{10} = &\frac{\pi[S_1(1-S_2) - CovS_{12}]}{\pi[S_1(1-S_2) - CovS_{12}] + (1-\pi)[(1-C_1)C_2 - CovC_{12}]}\\
&Y_{01}|\pi, S_{1,2}, C_{1,2}, CovS_{12}, CovC_{12}, N_{01} \sim Bino(N_{01}, p_{01})\\
\text{where } p_{01} = &\frac{\pi[(1-S_1)S_2 - CovS_{12}]}{\pi[(1-S_1)S_2 - CovS_{12}] + (1-\pi)[C_1(1-C_2) - CovC_{12}]}\\
&Y_{00}|\pi, S_{1,2}, C_{1,2}, CovS_{12}, CovC_{12}, N_{00} \sim Bino(N_{00}, p_{00})\\
\text{where } p_{00} = &\frac{\pi[(1-S_1)(1-S_2) + CovS_{12}]}{\pi[(1-S_1)(1-S_2) + CovS_{12}] + (1-\pi)[C_1C_2 + CovC_{12}]}
\end{aligned}
$$

Random Effects Model

$$S_j = P(T_J = 1|D = 1) = \int_{-\infty}^{\infty} P(T_{jk} = 1|D_k = 1, I_k = i_k)d\Phi(i_k) = \Phi\left(\frac{a_{j1}}{\sqrt{1 + b_{j1}^2}}\right), j = 1,2$$

$$C_j = P(T_J = 0|D = 0) = \Phi\left(\frac{a_{j0}}{\sqrt{1 + b_{j0}^2}}\right), j = 1,2$$

Let $Q_{k1} = \prod_{j=1}^{2}\Phi(a_{j1} + b_{j1}i_k)^{t_{jk}}[1 - \Phi(a_{j1} + b_{j1}i_k)]^{1-t_{jk}}$;

$Q_{k0} = \prod_{j=1}^{2}\Phi(a_{j0} + b_{j0}i_k)^{1-t_{jk}}[1 - \Phi(a_{j0} + b_{j0}i_k)]^{t_{jk}}$

$$
\begin{aligned}
L &\propto \prod_{k=1}^{N} P(T_{1k}=t_{1k}, T_{2k}=t_{2k}|\psi, I_k=i_k, D_k=d_k) \\
&= \prod_{k=1}^{N} \left\{ \pi \prod_{j=1}^{2} \Phi(a_{j1}+b_{j1}i_k)^{t_{jk}} [1-\Phi(a_{j1}+b_{j1}i_k)]^{1-t_{jk}} \right\}^{d_k} \\
&\times \prod_{k=1}^{N} \left\{ (1-\pi) \prod_{j=1}^{2} \Phi(a_{j0}+b_{j0}i_k)^{1-t_{jk}} [1-\Phi(a_{j0}+b_{j0}i_k)]^{t_{jk}} \right\}^{1-d_k} \\
&= \prod_{k=1}^{N} \{\pi Q_{k1}\}^{d_k} \{(1-\pi)Q_{k0}\}^{1-d_k}
\end{aligned}
$$

$\psi = (\pi, (a_{1d_k}, b_{1d_k}, a_{2d_k}, b_{21d_k})d_k = 0, 1)$ k=1,..,N

- Full Conditional Distributions of Parameters in the Random Effects Model

$$
\begin{aligned}
P(\pi|d_{1:N}, \alpha_\pi, \beta_\pi) \propto \pi^{\sum_{k=1}^{N} d_k + \alpha_\pi - 1} (1-\pi)^{N - \sum_{k=1}^{N} d_k + \beta_\pi - 1} \\
\implies P(\pi|d_{1:N}) \sim Beta\left(\sum_{k=1}^{N} d_k + \alpha_\pi, N - \sum_{k=1}^{N} d_k + \beta_\pi \right)
\end{aligned}
$$

$$
\begin{aligned}
P(d_k|t_{1k}, t_{2k}, \psi, i_k) \propto \{\pi Q_{k1}\}^{d_k} \{(1-\pi)Q_{k0}\}^{1-d_k} \\
\implies d_k|\psi, i_k \sim Bernoulli(p_k = \frac{\pi Q_{k1}}{\pi Q_{k1} + (1-\pi)Q_{k0}})
\end{aligned}
$$

$$
\begin{aligned}
P(a_{j1}, b_{j1}|t_{jk}, d_k, i_k) \propto \prod_{k=1}^{N} Q_{k1}^{d_k} d\Phi(a_{j1}, b_{j1}) \\
P(a_{j0}, b_{j0}|t_{jk}, d_k, i_k) \propto \prod_{k=1}^{N} Q_{k0}^{1-d_k} d\Phi(a_{j0}, b_{j0}) \\
P(i_k|t_{1k}, t_{2k}, d_k, a_{10,11,20,21}, b_{10,11,20,21}) \propto Q_{k1}^{d_k} Q_{k0}^{1-d_k}
\end{aligned}
$$

- A demo values for the data in Dendukuri & Joseph (2001)

```
n <- 10000
pi<- 0.85
s1 <- 0.25; s2 <- 0.8
c1 <- 0.95; c2 <- 0.7
covs <- 0.05; covc <- 0.0175

y1. <- n*pi*s1 ;  y.1 <- n*pi*s2
y0. <- n*pi*(1-s1) ; y.0 <- n*pi*(1-s2)

n1. <- n*pi*s1+n*(1-pi)*(1-c1) ; n.1 <- n*pi*s2+n*(1-pi)*(1-c2)
n0. <- n*pi*(1-s1)+n*(1-pi)*c1 ; n.0 <- n*pi*(1-s2)+n*(1-pi)*c2

(p111 <- s1*s2+covs)
## [1] 0.25
```

```
(p101 <- round(s1*(1-s2)-covs,4))
## [1] 0
(p011 <- (1-s1)*s2-covs)
## [1] 0.55
(p001 <- (1-s1)*(1-s2)+covs)
## [1] 0.2
(Y <- c(p111,p101,p011,p001)*n*pi)
## [1] 2125    0 4675 1700
sum(Y)
## [1] 8500

(p110 <- (1-c1)*(1-c2)+covc)
## [1] 0.0325
(p100 <- (1-c1)*c2-covc)
## [1] 0.0175
(p010 <- c1*(1-c2)-covc)
## [1] 0.2675
(p000 <- c1*c2+covc)
## [1] 0.6825

(N <- Y+c(p110,p100,p010,p000)*n*(1-pi))
## [1] 2173.75   26.25 5076.25 2723.75
sum(N)
## [1] 10000


y <- rbinom(n,1,pi)

#t1<- ifelse(y==1,rbinom(1,1,p111+p101),rbinom(1,1,p110+p100))
#t2<- ifelse(y==1,rbinom(1,1,p111+p011),rbinom(1,1,p110+p010))

#t1<- ifelse(y==1,rbinom(1,1,s1),rbinom(1,1,(1-c1)))
#t2<- ifelse(y==1,rbinom(1,1,s2),rbinom(1,1,(1-c2)))

t1 <- rbinom(n,1,s1)
t2 <- rbinom(n,1,s2)

data <- data.frame(cbind(y,t1,t2))
table(data$t1,data$t2,data$y)
## , ,  = 0
##
##
##        0    1
##   0  242  922
##   1   77  278
##
## , ,  = 1
##
##
##        0    1
##   0 1283 5104
##   1  392 1702
```

Stan

Gelman's homepage

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. The annals of applied statistics, 2(4), 1360-1383.

- A demo code for beta-binomial model

```
library(rstan)
# library(rstanarm)
stancode <- 'data {
  int<lower = 1> N;
  real<lower = 0> a;
  real<lower = 0> b;
}
transformed data { // these adhere to the conventions above
  real pi_ = beta_rng(a, b);
  int y = binomial_rng(N, pi_);
}
parameters {
  real<lower = 0, upper = 1> pi;
}
model {
  target += beta_lpdf(pi | a, b);
  target += binomial_lpmf(y | N, pi);
}
generated quantities { // these adhere to the conventions above
  int y_ = y;
  vector[1] pars_;
  int ranks_[1] = {pi > pi_};
  vector[N] log_lik;
  pars_[1] = pi_;
  for (n in 1:y) log_lik[n] = bernoulli_lpmf(1 | pi);
  for (n in (y + 1):N) log_lik[n] = bernoulli_lpmf(0 | pi);
}'

beta_binomial <- stan_model(model_code = stancode, verbose = TRUE)

# save(beta_binomial,file="/home/qs26/qushen26/stat2019_website/static/stat501/beta_binomial.stan")
```

```
load("beta_binomial.stan")
output <- sbc(beta_binomial, data = list(N = 100, a = 1, b = 1), M = 300, refresh = 0)
print(output)
plot(output, bins = 10)
```