Bayesian Hierarchical Modeling for Meta-Analysis

Overview

Meta-analysis is an important technique that combines information from different studies. When you have no prior information for thinking any particular study is different from another, you can treat Bayesian meta-analysis as a hierarchical model. This example illustrates two different applications of hierarchical modeling to a meta-analysis in medicine. The first application uses a normal approximation to the likelihood, and the second application uses the exact binomial likelihood.

The SAS source code for this example is available as a text file attachment. In Adobe Acrobat, right-click the icon in the margin and select Save Embedded File to Disk. You can also double-click the icon to open the file immediately.

Analysis

Consider the data from Yusuf et al. (1985), which summarize mortality after myocardial infarction from 22 studies. For each study, the data are in the form of $2 \times 2$ tables that consist of patients who are randomly assigned to receive beta-blockers or placebo. For study $j$, suppose that $trt_j$ is the number of deaths out of $trtN_j$ patients in the treatment group and $ctrl_j$ is the number of deaths out of $ctrlN_j$ patients in the control group.

The following statements read the data into SAS and create the Multistudies data set:

```sas
data multistudies;
  input studyid ctrl ctrlN trt trtN;
  datalines;
  1  3  39  3  38
  2 14  116  7 114
  3  11  93  5  69
  4 127 1520 102 1533
  5  27  365  28  355
  6  6  52  4  59
  7 152  939  98  945
  8 48  471  60  632
  9  37  282  25  278
 10 188 1921 138 1916
 11 52  583  64  873
 12 47  266  45  263
```

Suppose you want use a Bayesian random-effects model to estimate both the study-specific treatment effect and the pooled treatment effect. A Bayesian random-effects model assumes that there is no prior information for thinking that one study is different from the other. This assumption is also known as exchangeability.

Hierarchical Model Using Normal Approximation to the Likelihood

Normal approximation to binomial likelihood is a classical method that is commonly used in meta-analysis. Let \( Y_j \) be the odds ratio for the \( j \)th study:

\[
Y_j = \frac{\text{trt}_j(\text{ctrl}_N - \text{ctrl}_j)}{\text{ctrl}_j(\text{trt}_N - \text{trt}_j)}
\]

You can estimate the treatment effect \( \theta_j \), \( j = 1, \ldots, 22 \), through the means of an approximate normal distribution,

\[
\log(Y_j) \sim \text{normal}(\theta_j, s_j^2)
\]

where \( \theta_j \) is the study-specific effect and \( s_j^2 \) is the known variance of \( \log(Y_j) \). For \( s_j^2 \), because all studies have large sample sizes (more than 50 persons in each group in nearly all of the studies), you can use the approximate sampling variance of \( \theta_j \):

\[
s_j^2 = \frac{1}{\text{trt}_j} + \frac{1}{\text{trt}_N - \text{trt}_j} + \frac{1}{\text{ctrl}_j} + \frac{1}{\text{ctrl}_N - \text{ctrl}_j}
\]

If the odds ratios are exchangeable between the studies, then the treatment effect in each trial can be considered to be a random quantity drawn from some population distribution. In a Bayesian framework, this means that you can place a common prior distribution on the exchangeable random-effects parameters \( \theta_1, \ldots, \theta_{22} \),

\[
\theta_j \sim \text{normal}(\mu, \tau^2)
\]

where \( \mu \) is the population average of the treatment effect across all studies and \( \tau^2 \) is the between-study variation. Suppose the following noninformative priors are placed on the hyperparameters \( \mu \) and \( \tau^2 \):

\[
\mu \sim \text{normal}(0, sd = 3)
\]

\[
\tau^2 \sim \text{igamma}(shape = 0.01, scale = 0.01)
\]

The following DATA step creates the response variable, \( \log(Y_j) \), and its approximate sampling variance, \( s_j^2 \) for \( j = 1, \ldots, 22 \).
Hierarchical Model Using Normal Approximation to the Likelihood

data multistudies;
  set multistudies;
  logy=log(trt/(trtN-trt)/(ctrl/(ctrlN-ctrl)));
  sigma2=1/trt+1/(trtN-trt)+1/ctrl+1/(ctrlN-ctrl);
run;

The following MCMC procedure statements fit the described random-effects model:

proc mcmc data=multistudies outpost=nlout seed=276 nmc=50000 thin=5
  monitor=(OR Pooled);
  array OR[22];
  parms mu tau2;
  prior mu ~ normal(0, sd=3);
  prior tau2~ igamma(0.01,s=0.01);
  random theta ~n(mu, var=tau2) subject=studyid;
  OR[studyid]=exp(theta);
  Pooled=exp(mu);
  model logy ~ n(theta, var=sigma2);
run;

The PROC MCMC statement specifies the input data set as Multistudies and the output data set as Nlout, sets a seed for the random number generator, and requests a large simulation size with thinning. The MONITOR= option in the PROC MCMC statement outputs the analyses or saves the posterior samples for the specified variables. The ARRAY statement allocates an array variable OR of size 22, which is the number of random-effects parameters. The PARMS statement declares \( \mu \) and \( \tau^2 \) as model parameters, and the PRIOR statements specify their prior distributions.

The RANDOM statement declares \( \theta_j \) as random-effects parameters and specifies the normal prior distribution with hyperparameters \( \mu \) and \( \tau^2 \). The SUBJECT= option in the RANDOM statement specifies the group index as the data set variable studyid. The RANDOM statement generates 22 random-effects parameters, one for each unique value in the studyid variable. To monitor the function of 22 random-effects parameters, you must allocate an array and store the transformation of every random-effects parameter. The next programming statement, \( \text{OR[studyid]}=\exp(\text{theta}) \), does precisely this. As the procedure steps through the data set, it calculates the exponential of each \( \theta_j \) parameter and stores the result in the \( j \)th element of the OR array according to the studyid variable value (indexing). For this statement to work, the subject index variable studyid must have an integer value (1, 2, 3, and so on). The symbol Pooled calculates the exponential of the hyperparameter \( \mu \), which is the pooled treatment effect. The MODEL statement specifies the normal likelihood for the log odds ratios.

Figure 1 reports posterior summary statistics for each of the 22 odds ratios and for the Pooled estimate. The odds ratios provide the multiplicative change in the odds of mortality in the treatment group with respect to the control group. Some variation exists among the treatment effects. The pooled estimate of the odds ratio is 0.7849, indicating that mortality is less likely in the treatment groups.
You can use `%CATER` autocall macro to create a caterpillar plot of the odds ratios. The following statement takes the output from the `Nlout` data set and generates a caterpillar plot of all study-specific odds ratios and the pooled odds ratio:

```
%CATER(data=nlout, var=OR: Pooled);
```

Figure 2 shows the posterior means and the 95% equal-tail credible intervals of odds ratios from the 22 studies and for the pooled odds ratio. It looks like the treatment is most effective for Study 7 and least effective for Study 14. This result illustrates another advantage of the random-effects models: a fixed-effects model focuses on the coefficient effects after averaging study differences, whereas a random-effects model can help identify studies that are very different from others, such as ones with very high or low treatment effects.
Hierarchical Model Using Binomial Likelihood

The standard normal approximation method might not be appropriate because the approximation becomes less precise in the extreme probabilities (for example, 0,1). In the Multistudies data set, the mortality rates in the thirteenth study and nineteenth study are closest to 0 (≈ 0.03). Therefore, an alternative approach is to use the exact likelihood approach as opposed the normal approximation.

Emulating the binomial model as outlined in Spiegelhalter, Abrams, and Myles (2004), you can use the following model, where the treatment and control groups have their own binomial likelihood functions, with death probabilities $p_j$ and $q_j$ in the treatment group and control group, respectively:

$$
trt_j \sim \text{binomial}(trtN_j, p_j)
$$

$$
ctrl_j \sim \text{binomial}(ctrlN_j, q_j)
$$

Let the log odds for the control group be $\phi_j$ and let the log odds for the treatment group be $\theta_j + \phi_j$ as follows:

$$
\phi_j = \log(q_j/1-q_j)
$$

$$
\theta_j + \phi_j = \log(p_j/1-p_j)
$$
Then \( \theta_j \) is the log odds ratio:
\[
\theta_j = \log(p_j(1 - q_j)/(1 - p_j)q_j)
\]

If treatment group risks are exchangeable, then you can assume that \( \theta_1, \ldots, \theta_J \) are random-effects parameters that are drawn from some common prior distribution:
\[
\pi(\theta_j) = \text{normal}(\mu_\theta, \sigma_\theta^2)
\]

If the control group risks are also exchangeable, then you can consider \( \phi_1, \ldots, \phi_J \) as other random-effects parameters with the following common prior:
\[
\pi(\phi_j) = \text{normal}(\mu_\phi, \sigma_\phi^2)
\]

Suppose the following priors are placed on the model parameters:
\[
\begin{align*}
\pi(\mu_\theta) \cdot \pi(\mu_\phi) &= \text{normal}(0, sd = 3) \\
\pi(\sigma_\theta^2) \cdot \pi(\sigma_\phi^2) &= \text{igamma}(shape = 0.01, scale = 0.01)
\end{align*}
\]

The following PROC MCMC statements fit the described model:

``` SAS
proc mcmc data=multistudies outpost=blout seed=126 nmc=50000 thin=5
   monitor=(OR Pooled);
   array OR[22];
   parms mu_theta mu_phi s_theta s_phi;
   prior mu:~ normal(0, sd=3);
   prior s:~ igamma(0.01, s=0.01);
   random theta ~n(mu_theta, var=s_theta) subject=studyid;
   random phi ~n(mu_phi, var=s_phi) subject=studyid;
   p = logistic(theta + phi);
   q = logistic(phi);
   model trt ~ binomial(trtN, p);
   model ctrl ~ binomial(ctrlN, q);
   OR[studyid]=exp(theta);
   Pooled=exp(mu_theta);
run;
```

The PROC MCMC statement specifies the input data set as Multistudies and the output data set as Blout, sets a seed for the random number generator, and requests a large simulation size with thinning. The MONITOR= option in the PROC MCMC statement saves the posterior samples for the specified variables. The ARRAY statement allocates an array variable OR of size 22. The PARMS statement declares \( \mu_\theta, \mu_\phi, \sigma_\theta^2, \) and \( \sigma_\phi^2 \) as model parameters, and the PRIOR statements specify their prior distributions. The RANDOM statements declare \( \theta_j \) and \( \phi_j \) as random-effects parameters and specify a normal prior distribution for them. The SUBJECT= option in both RANDOM statements indicates the group index as data set variable studyid. The LOGISTIC functions perform logistic transformation from the log odds of the treatment and control groups to the corresponding binomial mortality probabilities. The MODEL statements specify the binomial likelihood for the treatment and the control groups. The programming statement in the next line,
Hierarchical Model Using Binomial Likelihood

OR[studyid]=exp(\theta), calculates the exponential of each ‘theta’ effect and stores the result in the OR array according to the studyid variable value (indexing). The symbol Pooled calculates the exponential of the hyperparameter \( \mu_\theta \), which is the pooled treatment effect.

Figure 3 reports posterior summary statistics for each of the 22 odds ratios.

**Figure 3** Posterior Summaries for the Binomial Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>N</th>
<th>Mean</th>
<th>Deviation</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR1</td>
<td>10000</td>
<td>0.7960</td>
<td>0.1410</td>
<td>0.7066</td>
<td>0.7867</td>
<td>0.8668</td>
</tr>
<tr>
<td>OR2</td>
<td>10000</td>
<td>0.7477</td>
<td>0.1221</td>
<td>0.6698</td>
<td>0.7428</td>
<td>0.8202</td>
</tr>
<tr>
<td>OR3</td>
<td>10000</td>
<td>0.7734</td>
<td>0.1308</td>
<td>0.6913</td>
<td>0.7639</td>
<td>0.8460</td>
</tr>
<tr>
<td>OR4</td>
<td>10000</td>
<td>0.7801</td>
<td>0.0817</td>
<td>0.7230</td>
<td>0.7758</td>
<td>0.8308</td>
</tr>
<tr>
<td>OR5</td>
<td>10000</td>
<td>0.8461</td>
<td>0.1322</td>
<td>0.7573</td>
<td>0.8274</td>
<td>0.9163</td>
</tr>
<tr>
<td>OR6</td>
<td>10000</td>
<td>0.7771</td>
<td>0.1345</td>
<td>0.6924</td>
<td>0.7656</td>
<td>0.8515</td>
</tr>
<tr>
<td>OR7</td>
<td>10000</td>
<td>0.6868</td>
<td>0.0762</td>
<td>0.6346</td>
<td>0.6859</td>
<td>0.7376</td>
</tr>
<tr>
<td>OR8</td>
<td>10000</td>
<td>0.8377</td>
<td>0.1113</td>
<td>0.7626</td>
<td>0.8286</td>
<td>0.9012</td>
</tr>
<tr>
<td>OR9</td>
<td>10000</td>
<td>0.7583</td>
<td>0.1104</td>
<td>0.6854</td>
<td>0.7533</td>
<td>0.8250</td>
</tr>
<tr>
<td>OR10</td>
<td>10000</td>
<td>0.7402</td>
<td>0.1104</td>
<td>0.6854</td>
<td>0.7533</td>
<td>0.8250</td>
</tr>
<tr>
<td>OR11</td>
<td>10000</td>
<td>0.7947</td>
<td>0.1113</td>
<td>0.7626</td>
<td>0.8286</td>
<td>0.9012</td>
</tr>
<tr>
<td>OR12</td>
<td>10000</td>
<td>0.8577</td>
<td>0.1288</td>
<td>0.7692</td>
<td>0.8414</td>
<td>0.9259</td>
</tr>
<tr>
<td>OR13</td>
<td>10000</td>
<td>0.7365</td>
<td>0.1201</td>
<td>0.6592</td>
<td>0.7334</td>
<td>0.8094</td>
</tr>
<tr>
<td>OR14</td>
<td>10000</td>
<td>0.9382</td>
<td>0.1443</td>
<td>0.8361</td>
<td>0.9168</td>
<td>1.0183</td>
</tr>
<tr>
<td>OR15</td>
<td>10000</td>
<td>0.7954</td>
<td>0.1174</td>
<td>0.7185</td>
<td>0.7849</td>
<td>0.8625</td>
</tr>
<tr>
<td>OR16</td>
<td>10000</td>
<td>0.8290</td>
<td>0.1265</td>
<td>0.7470</td>
<td>0.8138</td>
<td>0.8961</td>
</tr>
<tr>
<td>OR17</td>
<td>10000</td>
<td>0.8502</td>
<td>0.1424</td>
<td>0.7578</td>
<td>0.8290</td>
<td>0.9196</td>
</tr>
<tr>
<td>OR18</td>
<td>10000</td>
<td>0.8082</td>
<td>0.1391</td>
<td>0.7171</td>
<td>0.7908</td>
<td>0.8819</td>
</tr>
<tr>
<td>OR19</td>
<td>10000</td>
<td>0.7840</td>
<td>0.1338</td>
<td>0.6993</td>
<td>0.7748</td>
<td>0.8561</td>
</tr>
<tr>
<td>OR20</td>
<td>10000</td>
<td>0.8078</td>
<td>0.1185</td>
<td>0.7279</td>
<td>0.7948</td>
<td>0.8771</td>
</tr>
<tr>
<td>OR21</td>
<td>10000</td>
<td>0.7190</td>
<td>0.1049</td>
<td>0.6503</td>
<td>0.7182</td>
<td>0.7879</td>
</tr>
<tr>
<td>OR22</td>
<td>10000</td>
<td>0.7064</td>
<td>0.1077</td>
<td>0.6348</td>
<td>0.7055</td>
<td>0.7753</td>
</tr>
<tr>
<td>Pooled</td>
<td>10000</td>
<td>0.7810</td>
<td>0.0529</td>
<td>0.7446</td>
<td>0.7795</td>
<td>0.8151</td>
</tr>
</tbody>
</table>

You can use %CATER autocall macro to create a caterpillar plot of the odds ratios. The following statement takes the output from the Nlout data set and generates a caterpillar plot of all study-specific treatment effects and the overall treatment effect:

```
%CATER(data=blout, var=OR: Pooled);
```

Figure 4 shows 95% credible intervals of odds ratios from the 22 studies. Comparing Figure 2 and Figure 4, you can see that the posterior means and the credible intervals that are obtained by using the approximate normal likelihood and the binomial likelihood are very similar.
Figure 5 compares the kernel density plots of the odds ratios that are produced by using the approximate normal likelihood and the exact binomial likelihood. The first plot shows the comparison for the thirteenth study, and the second plot shows the comparison for the nineteenth study. Even though the observed mortality rates ($\approx 0.03$) are close to 0 for both studies, the kernel density plots that are produced by using the approximate normal likelihood and the exact binomial likelihood are very similar. Hence, this example shows very little sensitivity to the likelihood assumption. In general, this can be explained by the large sample sizes of the studies included in the meta-analysis.
**References**
