Lesson 11 Part A

Quiz, 8 questions

8/8 points (100%)

Congratulations! You passed!

Next Item

1 / 1
points

1.
Which of the following situations would call for a hierarchical model, due to the hierarchical structure of the data?

- You measure an individual's mood with a questionnaire for ten consecutive days.
- You survey a random sample of 100 individuals in your neighborhood who report their preferred produce market.
- You take blood pressure measurements from each of five individuals on ten consecutive days.

Correct
The blood pressure measurements are grouped within (also called "nested" within) individuals. Hence, you would expect two measurements from person A to be more similar than a measurement from person A and another from person B.

- You run an internet connection speed test on your computer each Monday morning for ten consecutive weeks.

---

1 / 1
points

2.
In hierarchical models, the observations are still conditionally independent, given their respective parameters. How then does such a model capture correlation among grouped observations?

- Grouped observations share common parameters, which themselves share a common distribution across groups.

Correct

The grouped observations are conditionally independent, given their common group parameters. However, if we integrate (marginalize) out this layer of group-specific parameters, leaving the hyperparameters and the observations only, the observations become dependent.

- Observation pairs from different groups require an extra parameter which induces negative correlation among them.
- Grouped observations always exhibit correlation, even in an independent model.
- Grouped observations require an extra parameter for each pair of observations which estimates the correlation among them.

---

1 / 1
points

3.
In previous lessons, we fit models to data representing percent growth in personnel for companies in two industries. Below are attached additional data from the original two industries (with 10 and six companies respectively), as well as three additional industries. Percent growth is reported for a total of 53 companies.

pctgrowth.csv

As usual, you can read the data into R with

```
dat = read.csv(file="pctgrowth.csv", header=TRUE)
```

where you will need to set R's working directory to where you have saved the file, or include the complete path to the file.

Rather than fit five separate models, one for each industry, we can fit a hierarchical model. As before, we assume a normal likelihood and common variance across all observations. Each industry will have it's own mean growth, and each of these means will come from a common distribution, from which we will estimate the overall mean and variability across industries.

Let i index the individual companies, and g_i indicate the industry (`grp` variable in `pctgrowth.csv`) for company i. Which of the following hierarchical models is the one described above?

$$
\begin{aligned}
& y_i \mid \theta_{g_i}, \sigma^2 \overset{\text{ind}}{\sim} \mathrm{N}(\theta_{g_i}, \sigma^2)\,, \quad i = 1, \ldots, 53, \quad g_i \in \{1, \ldots, 5\}, \\
& \theta_g \mid \mu, \tau^2 \overset{\text{iid}}{\sim} \mathrm{N}(\mu, \tau^2)\,, \quad g = 1, \ldots, 5, \\
& \mu \sim \mathrm{N}(0, 1\mathrm{e}6)\,, \\
& \tau^2 \sim \mathrm{IG}(1/2,\, 1 \cdot 3/2)\,, \\
& \sigma^2 \sim \mathrm{IG}(2/2,\, 2 \cdot 1/2)
\end{aligned}
$$

Correct

This model allows us to explore each of:

1) the mean growth for each industry (θ)

2) the overall mean growth across industries (μ)

3) the overall variability in mean growth across industries (τ^2)

4) the variability of company growth between companies within industries (σ^2).

All of these objectives are accomplished with a single model.

- [ ] $$
\begin{aligned}
& y_i \mid \theta_{g_i}, \sigma^2 \overset{\text{ind}}{\sim} \mathrm{N}(\theta_{g_i}, \sigma^2)\,, \quad i = 1, \ldots, 53, \quad g_i \in \{1, \ldots, 5\}, \\
& \theta_g \overset{\text{iid}}{\sim} \mathrm{N}(0, 1)\,, \quad g = 1, \ldots, 5, \\
& \sigma^2 \sim \mathrm{IG}(2/2,\, 2 \cdot 1/2)
\end{aligned}
$$

- [ ] $$
\begin{aligned}
& y_i \mid \theta_{g_i}, \sigma^2 \overset{\text{ind}}{\sim} \mathrm{N}(\theta_{g_i}, \sigma^2)\,, \quad i = 1, \ldots, 53, \quad g_i \in \{1, \ldots, 5\}, \\
& \theta_g \mid \tau^2 \overset{\text{iid}}{\sim} \mathrm{N}(0, \tau^2)\,, \quad g = 1, \ldots, 5, \\
& \tau^2 \sim \mathrm{IG}(1/2,\, 1 \cdot 3/2)\,, \\
& \sigma^2 \sim \mathrm{IG}(2/2,\, 2 \cdot 1/2)
\end{aligned}
$$

- [ ] $$
\begin{aligned}
& y_i \mid \theta_{g_i}, \sigma^2 \overset{\text{ind}}{\sim} \mathrm{N}(\theta_{g_i}, \sigma^2)\,, \quad i = 1, \ldots, 53, \quad g_i \in \{1, \ldots, 5\}, \\
& \theta_g \mid \mu \overset{\text{iid}}{\sim} \mathrm{N}(\mu, 1)\,, \quad g = 1, \ldots, 5, \\
& \mu \sim \mathrm{N}(0, 1\mathrm{e}6)\,, \\
& \sigma^2 \sim \mathrm{IG}(2/2,\, 2 \cdot 1/2)
\end{aligned}
$$

---

1 / 1
points

4.

Fit the hierarchical model from Question 3 in JAGS and obtain posterior mean estimates for each industry's mean growth (posterior mean for each θ_g).

Lesson 11 Part A

Quiz, 8 questions

8/8 points (100%)

We are interested in comparing these estimates to those obtained from a model that assumes no hierarchy (the ANOVA cell means model). We can approximate the posterior estimates for the five industry means under a noninformative prior by simply calculating the sample mean growth for the five industries. You can do this in R with:

```
means_anova = tapply(dat$y, INDEX=dat$grp, FUN=mean)
  ## dat is the data read from pctgrowth.csv
```

How do these compare with the estimates from the hierarchical model?

Hint: It might help to plot them with:

```
plot(means_anova)
points(means_theta, col="red") ## where means_theta are the posterior point
    estimates for the industry means.
```

- The estimates from the hierarchical model have **greater** variability than those from the ANOVA model, tending toward **larger** magnitudes.

- The estimates from the hierarchical model have **less** variability than those from the ANOVA model, tending toward **smaller** magnitudes.

Correct
This is a typical feature of hierarchical models, where estimates tend to "shrink" toward their mean in the next step of the hierarchy (in this case μ).

- The estimates from the hierarchical model have **greater** variability than those from the ANOVA model, tending toward **smaller** magnitudes.

- The estimates from the hierarchical model have **less** variability than those from the ANOVA model, tending toward **larger** magnitudes.

---

1 / 1
points

5.
In our hierarchical model for personnel growth, we assumed that the variability between companies within an industry was constant across industries (σ^2 was the same for all industries). Each of the following, except one, presents a reasonable approach to checking this model assumption. Which approach would be less informative than the others?

- Fit another model with separate σ^2 parameters for each industry, and calculate the posterior probability that they differ from each other by some specified amount.

-

Fit another model with separate σ^2 parameters for each industry, and compare the DIC values between the new and old models.

Lesson 11 Part A

Quiz, 8 questions

8/8 points (100%)

- Calculate the sample variance of growth within the industry for each of the five industries. Check if these variances are similar to each other.

- Calculate the posterior probability that $\sigma^2/\tau^2 > 1$ in the original model. If this probability exceeds a pre-determined amount, use a model with separate variance parameters.

Correct
While it is true that wildly different variances across industries will inflate a common σ^2, it is perfectly possible that the industry means are close to each other (low τ^2) while all industries exhibit similar variance among companies, each with large σ^2.

---

1 / 1
points

6.
Which of the following would yield the correct calculation of the observation level residual for Company i in the hierarchical model for personnel growth?

- $y_i - \theta_{g_i}$

Correct
We could obtain a distribution of the residual by calculating performing this calculation for each iteration of MCMC, or we could just get a posterior mean residual by using the posterior mean estimate of θ_{g_i}.

We can get an industry level residual by calculating $\theta_{g_i} - \mu$.

- $y_i - (\theta_{g_i} - \mu)$
- $\theta_{g_i} - \mu$
- $y_i - \mu$

---

1 / 1
points

7.

Suppose we are interested in making a prediction for the growth of an 11th (new) company from Industry 1 (`grp = 1`). Which of the following steps should we follow to simulate from the posterior predictive distribution for this new company's growth? Call it y^*.

- For each iteration of MCMC:

  Draw a sample y^* from a normal distribution with mean μ and variance $\tau^2 + \sigma^2$ using the current iteration of these parameters.

- For each iteration of MCMC:

  Draw a sample y^* from a normal distribution with mean μ and variance τ^2 using the current iteration of these parameters.

- For each iteration of MCMC:

  Draw a sample θ_1^* from a normal distribution with mean μ and variance τ^2. Then sample y^* from a normal distribution with mean θ_1^* and variance σ^2 using the current iteration of all parameters.

- For each iteration of MCMC:

  Draw a sample y^* from a normal distribution with mean θ_1 and variance σ^2 using the current iteration of these parameters.

Correct
Since the distribution of observations from Industry 1 only depends on θ_1 and variance σ^2, these are the only parameters we need to use to simulate growth for a new company from that industry.

---

1 / 1
points

8.
Suppose we are interested in making a prediction for the growth of a new company from a new industry which so far has not been observed. Which of the following steps should we follow to simulate from the posterior predictive distribution for this new company's growth? Call it y^*.

- For each iteration of MCMC:

  Draw a sample y^* from a normal distribution with mean μ and variance σ^2 using the current iteration of these parameters.

- For each iteration of MCMC:

Draw a sample θ^* from a normal distribution with mean μ and variance τ^2. Then sample y^* from a normal distribution with mean θ^* and variance σ^2, using the current iteration of all parameters.

Correct

Because this company is from a new industry, we must draw first from the posterior predictive distribution of industry means. Then once we have a mean for our new industry, we sample from the distribution of growth.

In a hierarchical normal model, this is equivalent to drawing a sample y^* from a normal distribution with mean μ and variance $\tau^2 + \sigma^2$ (see the supplementary reading on hierarchical normal models).

- For each iteration of MCMC:

  Draw a sample y^* from a normal distribution with mean μ and variance τ^2 using the current iteration of these parameters.

- For each iteration of MCMC:

  Draw a sample y^* from a normal distribution with mean θ_1 and variance σ^2 using the current iteration of these parameters.