05 - Hierarchical models

Matej Miočić

Introduction

We are interested in understanding whether there are any differences in measured VO_2 max values between boys and girls and between two training devices used in the study: cycle and treadmill. We analyzed the vo2max.csv dataset which contains 453 instances of the measured VO_2 max values of kids measured on a training device along with their age and sex. To answer our questions we used hierarchical normal model developed in stan.

Methods

First we removed the age feature since it had almost no correlation with the target variable. Then we split the dataset for each gender and device combination resulting in 4 subjects. We developed several models, to see if any of them performs better. Models named A have same variance across subjects, whereas models named B have subjects with their own variance. We used stan's default priors for all models.

First we used a simple normal model since the data look normally distributed and we model y_i as:

$$\textbf{Normal A}: y_i \sim normal(\mu_j, \sigma), \tag{1}$$

where each instance has a mean from their subject j and :

$$\textbf{Normal B}: y_i \sim normal(\mu_j, \sigma_j), \tag{2}$$

where each instance has a mean and a variance parameter from their subject j.

Second we used a gamma model since our target variable is a real positive number and we model y_i as:

$$\textbf{Gamma}: y_i \sim gamma(\alpha_j, \beta_j), \tag{3}$$

where each instance has a shape and a rate parameter from their subject j.

Third we used a hierachical normal model, since each subject is normally distributed and observations within each unit are correlated with one another. We model y_i as:

Hierarchical A :

$$\begin{aligned} \mu_j &\sim normal(\mu_\mu, \sigma_\mu) \\ y_i &\sim normal(\mu_j, \sigma), \end{aligned} \tag{4}$$

where each instance has a mean parameter from their subject j which is modeled from the same normal distribution with parameters μ_μ and σ_μ and:

Hierarchical B :

$$\begin{aligned} \mu_j &\sim normal(\mu_\mu, \sigma_\mu) \\ \sigma_j &\sim normal(\mu_\sigma, \sigma_\sigma) \\ y_i &\sim normal(\mu_j, \sigma_j), \end{aligned} \tag{5}$$

where each instance has a mean parameter from their subject j which is modeled from the same normal distribution with parameters μ_μ and σ_μ and a variance parameter from their subject j which is modeled from the same normal distribution with parameters μ_σ and σ_σ.

For every model we checked the diagnostics on our fits and they gave us no reasons for concern.

We decided to use Hierarchical A model for further analysis, since a hierarchical model is suitable for our dataset where observations within each subject are correlated with one another (we can assume participants used the same assigned device and children exhibit similar athletic aptitudes). The posteriror checks were similar between the models so we choose the simpler of the 2, where each subject has the same variance.

Then we answered our questions by looking at the obtained means and variances of our draws for each subject. We also calculated the probability that each subject has a higher VO_2 max value than the population mean by using the r function **pnorm** for each draw.

Results

We started the analysis by checking whether age has any linear correlation (Figure 1) with our target variable VO_2 max value.

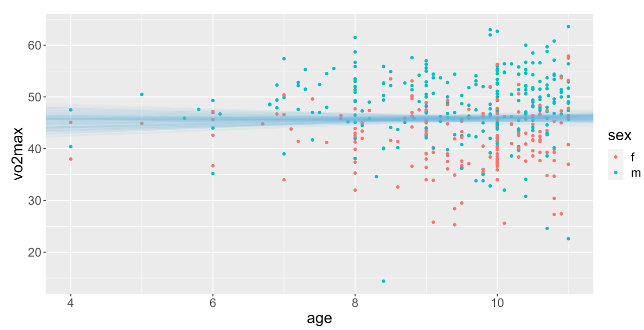

Figure 1. Fitted linear regression lines from 100 draws using a simple linear regression model on whole population using only age. For extra information about dataset male kids are coloured in blue and female kids in red.

We observe that on average the fitted line is almost horizontal, which means the slope is approximately 0. This means that VO_2 max value is not linearly dependent on age in our dataset.

We then proceeded to fit our dataset which was split into 4 subjects to all 5 models described in methods section. Since we're interested in differences on subject level, we compare their means with 90% HDI between all models and our sample mean (Figure 2).

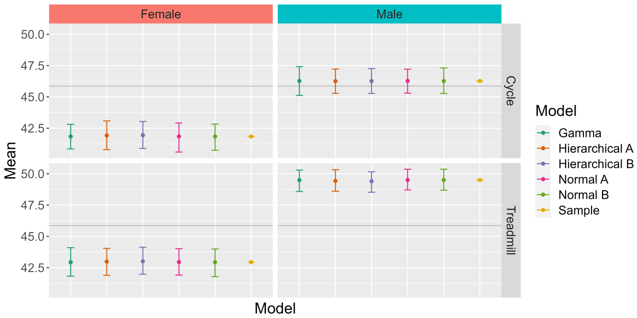

Figure 2. Mean estimates for each subject for each model with 90% HDI intervals, sample means and population mean (the gray horizontal line).

We also performed posterior checks on mentioned models, but we show only posterior check for hierarchical A model (Figure 3), since they were all similar and we have space limitation on our report.

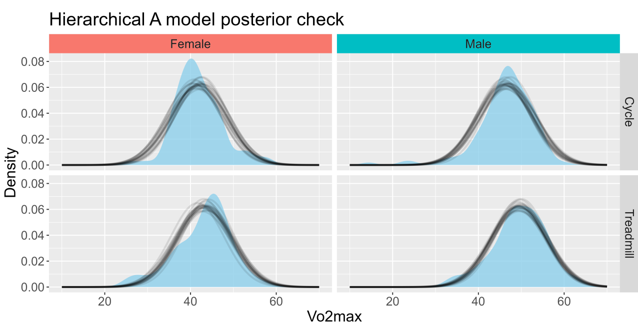

Figure 3. Posterior check for hierachical A model for each subject.

We can observe that model approximately captures the shape for all 4 subjects.

Table 1. Mean value of μ_j of VO_2 max of all draws for each subject with its standard deviation.

Female	Male	
41.9 ± 0.70	46.3 ± 0.59	Cycle
43.0 ± 0.65	49.4 ± 0.52	Treadmill

In Table 1 we show mean values of μ_j which were obtained from normal distribution with parameters: $\mu_\mu = 45.0 \pm 4.38$ and $\sigma_\mu = 7.26 \pm 8.14$. All subjects had the same variance which is: $\sigma = 6.37 \pm 0.21$.

We can see that the mean value of μ_j of VO_2 max is highest for participants who are male and were tested on a treadmill. The lowest mean was for participants who are female and were tested on a cycle. We can also observe differences when observing only sex and only devices separately. When observing sex, we see that male participants have higher VO_2 max value on average. When observing device, we can observe higher mean when participants were using a treadmill.

In Figure 4 we show an example how we calculated the percentage that a subject has a higher VO_2 value than the population mean = 45.86 ± 7.04. To calculate the final percentage we do the same for every draw and take the mean percentage.

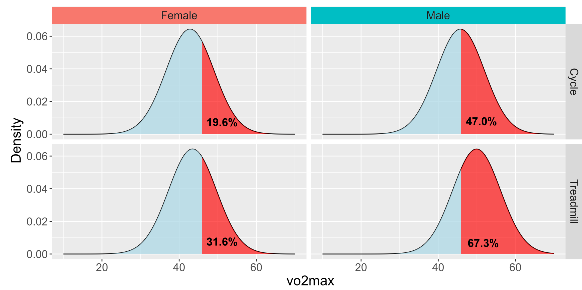

Figure 4. An example of proportion of density of the first draw where VO_2 max value is higher than population the mean for each subject (coloured in red).

Table 2. Mean probability with standard errors for each subject that its VO_2 value is higher than the population mean.

Female	Male	
$26.9\% \pm 3.7\%$	$52.4\% \pm 3.7\%$	Cycle
$32.7\ \% \pm 3.7\%$	$71.1\% \pm 2.9\%$	Treadmill

In Table 2 we show the differences between all subjects in terms of probability that their VO_2 is higher than population mean. We observe high chance when the participant is male and is using a treadmill.

With the same approach we can also calculate the percentage for sex and training device separately. We can just combine draws from subject 1 (Female + Cycle) and subject 3 (Female and Treadmill) to get all female draws.

Table 3. Mean probability with standard errors for each possible category that its VO_2 max value is higher than the population mean.

Female	$29.8\% \pm 4.7\%$
Male	$61.8\% \pm 9.9\%$
Cycle	$40.0\ \% \pm 13.2\%$
Treadmill	$51.9\% \pm 19.5\%$

In Table 3 we further show the impact on VO_2 max values based on sex and training device. We can see that being a male has approximately 32 increase in percentage points when compared to female sex of having its VO_2 max value higher than population mean. Whereas doing the experiment on a treadmill has approximately 12 percentage points increase over doing the experiment on a cycle.

Conclusion

We analysed whether there are any differences in measured VO_2 max values between boys and girls and between two training devices used in the study. We found out that being a male increases your odds of having a higher VO_2 max value the most. Using a treadmill increases your odds more as opposed to using a cycle, but more if you are a male.