Chapter I

Introduction

Expanding upon our preliminary exploration (Homework 1) of the impact of fertilizers on wheat yield, in this chapter we apply a hierarchical modeling approach to deepen our understanding. The primary objective is to compare the results obtained through hierarchical modeling with those from Homework 1. We will try to answer the same three questions:

Q1: How do fertilisers fare against each other on average?

Q2: What is the probability for each fertiliser that it will on average give the largest yield (larger than the other two)?

Q3: What is the probability that a particular fertiliser will give the largest yield next year?

Model

We applied a consistent methodological framework from our initial project, but instead of using three independent normal models for each fertilizer, we applied a hierarchical model with the default priors in Stan, defined as:

$$y_{s,i}|\mu_s,\sigma_s \sim \mathcal{N}(\mu_s,\sigma_s), \forall s \in \{1, 2, 3\} \quad (1)$$
$$\mu_s|\mu_\mu,\sigma_\mu \sim \mathcal{N}(\mu_\mu,\sigma_\mu) \quad (2)$$
$$\sigma_s|\mu_\sigma,\sigma_\sigma \sim \mathcal{N}(\mu_\sigma,\sigma_\sigma) \quad (3)$$

Within the hierarchical model, two links were established - linking the subject level means, μ_s, and the subject level standard deviations, σ_s, with normal distributions in equations (2) and (3), respectively.

Results

MCMC diagnostics, including Gelman-Rubin (R-hat), trace plots, and effective sample size, all indicated satisfactory results, affirming that MCMC chains have effectively explored parameter space and converged to the posterior distributions. A comparison between the hierarchical and non-hierarchical normal model in a form of group level and subject level means estimates, is shown in Figure 1. Additionally, the results are compared in Table 1, providing answers to the questions Q1, Q2, and Q3 from the introduction.

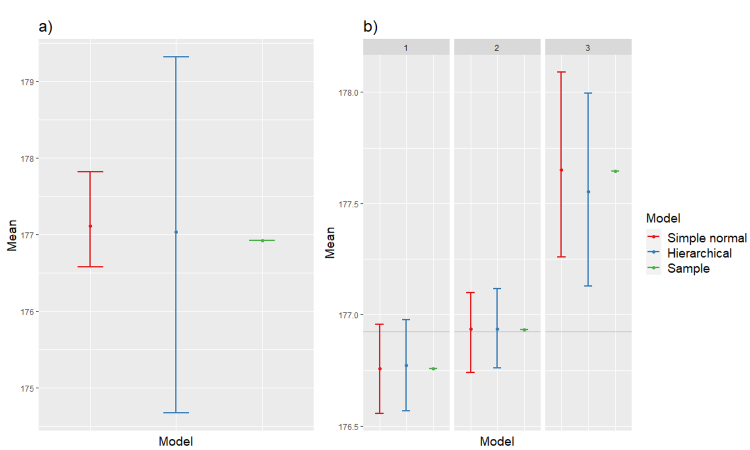

Figure 1. Comparison of the means and 90% HDI between both models and our sample, where: a) shows group level mean (for all fertilizers combined) and b) shows subject level means (for each fertilizer separately).

Table 1. Comparison of results obtained with simple normal and hierarchical model, where results are expressed as probability ± Monte Carlo standard error.

Q1: Pairwise comparison of fertilizers average yields		
	Simple normal model	Hierarchical model
$\mu_2 > \mu_1$	$(85.53 \pm 0.66)\%$	$(83.98 \pm 0.78)\%$
$\mu_3 > \mu_1$	$(99.68 \pm 0.17)\%$	$(99.15 \pm 0.22)\%$
$\mu_3 > \mu_2$	$(99.10 \pm 0.31)\%$	$(98.43 \pm 0.27)\%$
Q2: Probability of a fertiliser giving the largest yield on avg.		
	Simple normal model	Hierarchical model
fertilizer 1	$(0.10 \pm 0.05)\%$	$(0.48 \pm 0.15)\%$
fertilizer 2	$(0.83 \pm 0.22)\%$	$(1.43 \pm 0.26)\%$
fertilizer 3	$(99.07 \pm 0.31)\%$	$(98.09 \pm 0.36)\%$
Q3: Probability of a fertiliser giving the largest yield next year		
	Simple normal model	Hierarchical model
fertilizer 1	$(14.40 \pm 0.35)\%$	$(15.00 \pm 0.36)\%$
fertilizer 2	$(17.32 \pm 0.37)\%$	$(19.32 \pm 0.39)\%$
fertilizer 3	$(68.28 \pm 0.46)\%$	$(65.68 \pm 0.48)\%$

In Figure 1.a), we notice that hierarchical model estimate for the group level means is closer to the sample mean, but also has higher uncertainty, than non-hierarchical normal model estimate. This higher uncertainty comes from two sources: uncertainty in the subject level means and uncertainty in the group level mean. In Figure 1.b) we notice that mean estimates on subject level are closer to the sample mean of all fertilizers combined (gray line), due to Bayesian shrinkage effect. This effect is particularly noticeable in the case of fertilizer 3, which had the greatest sample mean and deviated significantly from the sample means of the other two fertilizers.

Conclusion

From Table 1, we observe that the hierarchical model gives similar answers to questions Q1, Q2, and Q3 as in Homework 1. Fertilizer 3 is still the one with the best performance, however, with the hierarchical model, the probabilities associated with this fertilizer in all three questions are slightly lower compared to simple normal model. In contrast, the probabilities for fertilizers 1 and 2 are slightly higher. This is the result of the model's information pooling in hierarchical modeling, where information about one subject influences inferences for all other subjects.

Chapter II

Introduction

VO_2 max is the maximum rate of oxygen consumption attainable during physical exertion. In this chapter, we analyze how VO_2 max changes with age and how these trends differ between boys and girls and between two different training devices. This analysis will give us a better understanding of the impact of age and gender on the physical performance of school-aged kids.

Data

In this project we used *vo2max.csv* dataset, that contains the data about physical performance of kids, measured on two different training devices, along with the age of kids for a certain measurement. The dataset reports relative VO_2 max measure, which is the maximum rate of oxygen consumption attainable during physical exertion normalised by weight, expressed as milliliters of oxygen per kilogram of body weight per minute (ml/kg/min). VO_2 *max* is our target variable. The other available features are:

- *age*: 4, 5, 6, 7, 8, 9, 10 or 11.
- *sex*: male or female.
- *device*: cycle or treadmill.

Model

The dataset was divided into four groups based on gender and the device used for measurement, resulting in the following groups:*male treadmill*, *male cycle*, *female treadmill*, and *female cycle*. For each group, linear regression was applied using age as the independent variable, default Stan priors, and a hierarchical structure established with two links:

$$y_{s,i}|\alpha_s, \beta_{s,i}, \sigma, i \sim \mathcal{N}(\alpha_s + \beta_s i, \sigma) \quad (4)$$

$$\alpha_s|\mu_\alpha, \sigma_\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha) \quad (5)$$

$$\beta_s|\mu_\beta, \sigma_\beta \sim \mathcal{N}(\mu_\beta, \sigma_\beta) \quad (6)$$

Initially, during the fitting process, a warning was encountered indicating that 12% of transitions ended with divergence. To address this, we set the *adapt_delta* to 0.99 and *max_treedepth* to 20, leading to a notable reduction in divergences, now occurring in only 1% of the samples.

Results

MCMC diagnostics, including the trace plots, assessments of R-hat values and effective sample size, all indicated satisfactory results, affirming that MCMC chains have successfully converged to posterior distributions.

The Table 2 presents the average values, along with their uncertainties, for the intercept and slope of the linear model across all four groups. Figure 2 illustrates that the estimated means at both the group and subject levels align closely with the samples mean, suggesting that our model effectively captures the underlying data patterns.

Table 2. Average linear regression coefficient values and their uncertainty for all four groups.

Group	Intercept (α_s)	Slope (β_s)
male treadmill	44.64 ±0.09	0.55 ±0.01
male cycle	43.12 ±0.10	0.34 ±0.01
female treadmill	42.08 ±0.10	0.10 ±0.01
female cycle	40.86 ±0.13	0.11 ±0.01

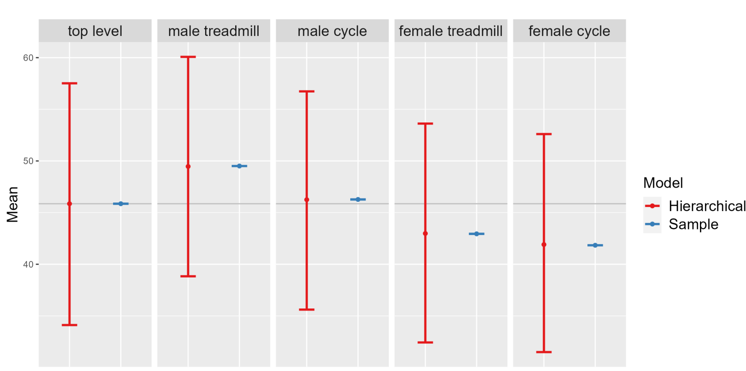

Figure 2. Estimated means and 90% HDI at the top (group) level and the subject level.

In Table 2, all mean beta coefficients are positive, indicating that VO_2 max increases with age in each of the four groups. The steepest rise with age is observed in the male treadmill group, with a $\beta = 0.55 \pm 0.01$, followed by the male cycle group with $\beta = 0.34 \pm 0.01$. Given that these two groups also have the highest intercepts, we can observe that boys, in general, tend to have higher VO_2 max than girls. A similar conclusion can be drawn from Figure 3, which illustrates the probability for each group to have the highest and lowest VO_2 max.

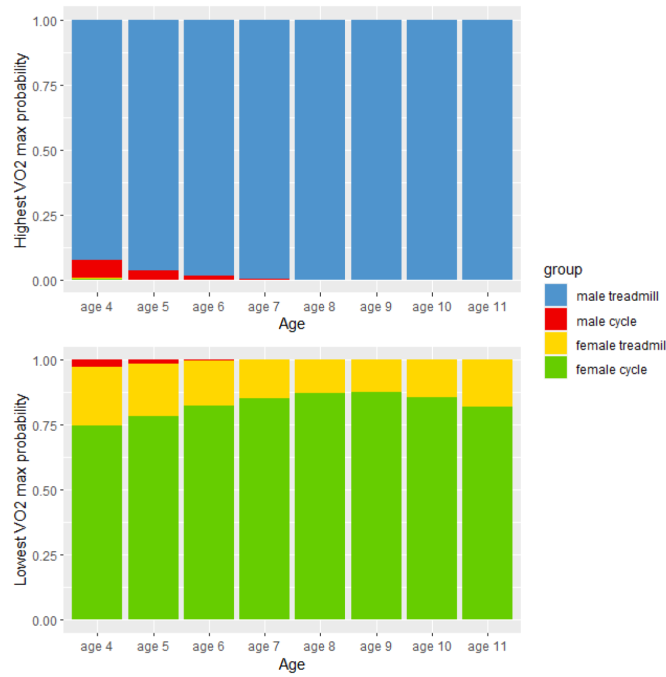

Figure 3. Probabilities for each group to achieve the highest and lowest VO_2 max values among all groups, categorized by different age groups. The upper image shows probabilities for the highest VO_2 max, while the lower image shows probabilities for the lowest VO_2 max.

In Figure 3, across all ages, the male treadmill group has the highest probability of achieving the maximum VO_2 max, followed by the male cycle group. However, after the age of 8, the male treadmill group has nearly a 100% probability of having the highest VO_2 max due to its larger beta coefficient compared to the male cycle. The lower image in Figure 3 shows that the female cycle group has the highest probability of having the lowest VO_2 max, followed by the female treadmill group. An interesting trend change is observed between age 9 and 10, where the probability of the female treadmill group having the lowest VO_2 max starts to increase at the expense of the female cycle group. This shift is attributed to the slightly lower beta coefficient ($\beta = 0.10 \pm 0.01$) for the female treadmill compared to the female cycle ($\beta = 0.11 \pm 0.01$).

Conclusion

The analysis of the obtained results provided insights into how the values of VO_2 max are influenced by age, gender, and the device used during measurement. VO_2 max values increase with age in children between 4 and 11 years, which is expected due to significant physical growth and development during this period. As children grow, their bodies become more efficient in utilizing oxygen. Additionally, it was observed that male children have higher VO_2 max than females across all age groups and on both devices: cycle and treadmill. Furthermore, we noticed that the treadmill device leads to a higher rate of oxygen consumption compared to the cycle. Thus, the highest VO_2 max across all age groups was observed in the male treadmill group, while the lowest was in the female cycle group.