

# MS Comprehensive Exam 2022

STA 207 (100 points)

Read the instructions on Canvas carefully!

Name:

Student ID:

```
library(lme4)
```

In this exam, we investigate the `ChickWeight` dataset in R. You can load the data using the following commands. Carefully read the help file of `ChickWeight` before working on the following questions.

```
data(ChickWeight)
```

(a) Briefly summarize all variables in the data set. You need to provide the definition of the variable and quantitative summary.

Solution: (Type your answer here)

```
# (a). (Type your code in the space below, if any)
```

(b) Visualize the weights of each chicks over time in one plot, where (i) each chick is represented by one solid curve, and (ii) the diet is color-coded as black (1), red (2), green (3), and blue (4). In addition to the required visualization, you may add any supporting curves, symbols, or any additional plots that you find informative.

Solution: (Type your answer here)

```
# (b). (Type your code in the space below, if any)
```

(c) Write down an appropriate one-way ANOVA model to answer the question whether there is any changes in mean weights at Day 20 across the four diet group. To receive full credits, you need to (i) write down the model, explain your notation, constraint(s) and/or assumptions; (ii) state the null and alternative hypotheses; (iii) state the test result. You can find basic LaTeX commands at the end of this file.

Solution: (Type your answer here)

```
# (c). (Type your code in the space below, if any)
```

(d) For the model fitted in (c), carry out necessary diagnostics to check if the model assumptions are valid. What are your findings?

Solution: (Type your answer here)

```
# (d). (Type your code in the space below, if any)
```

(e) Write down an appropriate two-way ANOVA model with fixed effect to answer the question whether there is any differences in growth rates across the four diet groups. Here the growth rate can be roughly seen as the effects of Time on weight. To receive full credits, you need to (i) write down the model, explain your notation, constraint(s) and/or assumptions; (ii) state the null and alternative hypotheses; (iii) state the test result. Hint: You may want to recycle the answer in (c) to save time.

Solution: (Type your answer here)

*# (e). (Type your code in the space below, if any)*

---

(f) We want to take the chick-specific effect into account. The new mixed effect model is based on the model in (e), where Time is treated as a continuous covariate instead of a categorical factor, and a random intercept and a random slope (of Time) are added into the model. Report the fitted coefficients of the fixed effects, and summarize your findings from this model. Hint: You do not need to write down the new model, but you may find it helpful.

Solution: (Type your answer here)

*# (f). (Type your code in the space below, if any)*

---

(g) Assume that the chicks in each diet are randomly selected from the same population, i.e., the enrollment of chicks is independent from any other factors. State the Stable Unit Treatment Value Assumption, write down the potential outcomes (weight at Day 20), and verify whether the randomization assumption holds. (This question will be replaced by another, since causal inference will not be covered this quarter.)

Solution: (Type your answer here)

---

LaTeX commands

$$Y_{i,j,k}, \mu, \alpha, \beta, \epsilon, i, j, H_0, H_a, \neq, =, \dots$$