STAT 213 (SUMMER 2021): HW4 (SLOS A2, A3, C2)

DUE ELECTRONICALLY BY THE CLASS TIME, FRIDAY 7/02/21

Instructions and Technical Tips

Write up your solutions and save/upload the file(s) to the RStudioPro server in the folder `stat213/turnin/hw5/` inside your Home directory. Include `hw5` (all lowercase) in your filenames.

Using an RMarkdown document is recommended, but not required.

RMarkdown Format. If you do use RMarkdown, R code should go in code chunks, and verbal commentary (and any math equations) should go outside code chunks.

Use section headings (a line starting with one or more `#` symbols) to demarcate the start of a problem. You can nest headings by using an additional `#` symbol for each level of nesting: one for a top-level heading, two for the next level within that, etc.

Periodically "Knit" your file to verify that it is working correctly.

"Knitting" Troubleshooting. If your code runs chunk by chunk but won't Knit, try clearing your environment (broom icon in the upper right) and running chunk by chunk from the start again. The most common cause is an undefined variable. This can happen if you change your variable names some places but not others, but a very common reason is that you read in the data from a file using a menu button instead of using the `read.file()` command, which means the "reading in the data" step is not recorded in your document. If you can't Knit, I won't be able to run your code either!

If you aren't able to Knit directly to `.pdf`, it may be because you are using special characters (such as $\neq$) in your `.Rmd`. Replace these with plain text and try again.

If you have done this and are still unable to Knit directly to `.pdf`, convert your Knitted `.html` or `.docx` into a `.pdf` (possibly by "printing" the `html` from your browser to a file) and save that there. A uniform file format across students will make grading 64 problem sets much more streamlined. Thanks!

What to turn in. If using Markdown, turn in the source file (ending in `.Rmd`) and the Knitted output file (preferably as `.pdf`). Otherwise, just turn in a `.pdf`.

Date: June 25, 2021.

Verifying receipt of your work. When I run the script that collects your work after the due date, you will see a "receipt" file in the `~/stat213/receipts/hw5/` directory. If this does not appear within 24 hours or so of you submitting your assignment, let me know.

Problems

1. (SLO C2) **Caterpillar metabolism revisited.** Use the model you fit in problem 2 on HW4 that used (log) body size (`LogBodySize`) to predict (log) metabolic rate (`LogMrate`) in caterpillars.

   The linear model was

   $$\widehat{\texttt{LogMrate}}_i = 1.3066 + 0.9164 \cdot \texttt{LogBodySize}_i$$

   The R code used to fit this model was:

   ```
   library(Stat2Data)
   data(MetabolicRate)
   LogLogModel <- lm(LogMrate ~ LogBodySize, data = MetabolicRate)
   ```

   Use this data and model to answer the following:

   (b) Find a 95% **confidence interval** for the **mean metabolic rate of a subpopulation of manduca sexta caterpillars weighing 1g**, and a 95% **prediction interval** for the metabolic rate of a **particular caterpillar weighing 1 gram**. Interpret both intervals in context. (Hint: Calculate both intervals on the log scale first, and then transform each endpoint back to the original scale)

   (c) Repeat part (b) for a caterpillar weighing 10 grams.

2. (SLOs A2, A3, D1) **Caterpillar metabolism by growth period.** Caterpillars go through free growth periods during each stage of their life. However, these periods end as the animal prepares to molt and then moves into the next stage of life. A biologist is interested in checking to see whether two different regression lines are needed to model the relationship between metabolic rates and body size of caterpillars for free growth and no free growth periods.

   (a) Identify the multiple regression model for predicting log metabolic rate `LogMrate` from log body size `LogBodySize` and an **indicator variable** for free growth: `Ifgp` = 1 for observations recorded during periods of free growth, `Ifgp` = 0 otherwise. The model should allow for two different regression lines (slope and/or intercept) depending on free growth status. (You will not actually fit this model in this problem; just write down the regression equation in symbols)

   (b) Identify the multiple regression model for predicting `Mrate` from `BodySize` and `Ifgp` for which the **rate of change** in predicted `Mrate` with respect to `BodySize` is the same during free growth and no free growth periods. That is, the model allows for faster (or slower) growth in general during free growth periods, and allows `BodySize` to affect the predicted growth rate, but each additional unit of `BodySize` has the same influence on predicted growth rate whether the caterpillar is in a free growth period or not.

   (c) Identify the **full and reduced models** that would be used in a **nested F-test** to check whether **one or two regression lines are needed** to model metabolic rates. Assume that if separate regression models are needed, they can have different intercepts and different slopes.