

Analyzing discrete data with GLMs

Homework assignment

DAM

04 April 2018

Notes on formatting your homework submission:

- No template this week. This is a simple enough assignment that you should be able to initiate your own `.Rmd` file.
- `R markdown` interprets hash tags/octothorpes/pound signs as headings. Use them to structure your document in the same order as the Assignment.
- Include your responses as text outside the R code chunks.
- For this assignment, include your code as an Appendix using the appropriate chunk from the mini-report assignment.
- Save as `.pdf` before uploading to Blackboard (or for \LaTeX users, `knit` straight to `.pdf`)

Data structure

We'll use the `starwars` data from the `tidyverse` package again. Load it now and review the data. *Note that as an object from the `tidyverse` the data are stored differently than other R objects. Do not use the `str()` command. Simply call the `starwars` object by its name.*

A basic pairwise comparison

Use a graph and statistics to answer a basic question: Are there more female or male characters in this dataset?

1. Subset the data to include just males and females
2. Present an appropriate `ggplot` to show the number of males and females. Give a descriptive caption.
3. Identify a relevant consideration about these data. Identify and justify an appropriate statistical test.
4. Perform this test. Answer the basic question above, supported by relevant test statistics.