

STAT 7630: Homework 8
(Due: Tuesday, 12/03/2024)

Note: Show all your work for the necessary steps to receive full credit.

Please turn in the HW on paper, hand-written and/or typed. For computational problems, return only the relevant parts of the output with comments/annotations. Questions taken from the textbook are marked with BR for “Bayes Rules!”. From the code you are using to answer the problems, turn in the relevant output, and the figures (if requested), preferably printed from the output. No need to turn in your code or long lists of generated samples.

Please disclose any use of AI in your solutions. Regardless of whether you use it or not, make sure you submit your own work, not copy from other source(s). Any suspicion of AI use will result in automatic 0 or substantial point loss in any question.

BR Chapter 12:

Do Exercises 12.3 and 12.4.

BR Chapter 13:

Do Exercises 13.1, 13.4 (b,c,d), 13.5, and 13.7.

Hints: In 13.7 (a), the `pp_check` step is optional for this part.

In 13.7 (b), report the estimated model on the probability scale.

For 13.7, below is some possibly helpful R code for reading in the `hotel_bookings` dataset. If you are using `stan_glm` from the `rstanarm` package, most of this is unnecessary as you can directly use the variables in the `hotel_bookings` data frame.

```
library(bayesrules)
data(hotel_bookings)
hotel_bookings
y <- hotel_bookings$is_canceled
y <- as.numeric(y) - 1  # Subtracting 1 converts the 1's and 2's to 0's and 1's.
x1 <- hotel_bookings$lead_time
x2 <- hotel_bookings$previous_cancellations
x3 <- hotel_bookings$is_repeated_guest
x4 <- hotel_bookings$average_daily_rate
```

Additional Questions (AQs):

AQ1. Read Exercise 12.5 in the *Bayes Rules!* book to learn about the `bald_eagles` data.

- (a) Fit a Poisson regression model relating Y to X_1 and X_2 . Use vague/uninformative priors.
- (b) Comment on whether X_1 and/or X_2 are apparently important or unimportant in explaining the count of eagle sightings.
- (c) In whatever way you choose, assess the fit of the Poisson model and comment on it.

(d) Fit a model that includes X_1 , X_2 , and the interaction $X_1 \times X_2$. Use any model selection technique of your choice to decide between the interaction model and the no-interaction model.

Below is some possibly helpful **R** code for reading in the **bald_eagles** data set. If you are using **stan_glm** in the **rstanarm** package, much of this is not necessary—you can directly use the variables in the **bald_eagles** data frame and use **year:hours** to specify the interaction.

```
# UPDATE: If you are using base R and NOT 'stan_glm' to fit the Poisson regression,
# then it seems like you encounter numerical issues when fitting the interaction model.
# To correct this (only necessary if using base R), redefine the 'Year' predictor (x1)
# to be 'Years since 2000'. This cuts down on the huge values of x1 and x1x2.

# Modified code:
x1 <- bald_eagles$Year - 2000 # Years since 2000
x2 <- bald_eagles$Hours
y <- bald_eagles$Sightings
```