

# Homework 9 – Stat 230 – Fall 2022

**Due date: Friday, November 11**

Complete the following exercises and submit your assignment via gradescope (linked on the course webpage).

## Note

When the book asks you to “Interpret the model in terms of the odds ratio,” it is asking you to interpret the slope of the logistic regression model.

**Problems to start after class Nov 4**

### Q1

Chapter 8, exercise E.1 parts (a) and (c)

### Q2

Read the prompt for Chapter 8 exercise E.7 , but then answer the following questions.

```
skincancer <- read.csv("https://aloy.rbind.io/kuiper_data/SkinCancer.csv")
```

- (a) Fit a Poisson regression model that uses the continuous version of age (`Age.Midpoint`) to describe the non-melanoma skin cancer incidence. Don’t forget to include the population as an exposure offset in your model. Report the fitted model equation.
- (b) Interpret the slope for age in context.
- (c) What is the predicted cancer rate for the 55-64 ages group (where `Age.Midpoint` is 60)?
- (d) Create a scatterplot of the incidence rate against the continuous version of age (`Age.Midpoint`). Do you have any concerns about using your Poisson regression model from part (a)?

## Problems to start after class Nov 7

### Q3

Chapter 8, exercise E.7

To load the data and extract only the rows for Minneapolis-St. Paul, run the following code:

```
msp <- read.csv("https://aloy.rbind.io/kuiper_data/SkinCancer.csv") |>
  dplyr::filter(Town == "Minneapolis-St Paul, Minnesota")
```

### Q4

Chapter 8, exercise E.8

To load the data and extract only the rows for Minneapolis-St. Paul, run the following code:

```
dfw <- read.csv("https://aloy.rbind.io/kuiper_data/SkinCancer.csv") |>
  dplyr::filter(Town == "Dallas-Fort Worth, Texas")
```

## Problems to start after class Nov 9

### Q5

Chapter 8, exercise E.5

We'll talk about overdispersion and zero-inflation on Friday, so for part (f) focus on other ways to check the model's adequacy.

```
apexams <- read.csv("https://aloy.rbind.io/kuiper_data/APEXams.csv")
```

### Q6

Chapter 8, exercise E.6

```
hurricanes <- read.csv("https://aloy.rbind.io/kuiper_data/Hurricanes.csv")
```

In part (a), you can calculate the response variable, the total number of hurricanes, using `dplyr::mutate()`:

```
hurricanes <- hurricanes |> dplyr::mutate(count = cat1 + cat2 + cat3 + cat4 + cat5)
```

In part (b), you can calculate the response variable, the total number of intense hurricanes, using `dplyr::mutate()`:

```
hurricanes <- hurricanes |> dplyr::mutate(intense = cat3 + cat4 + cat5)
```