# Quiz 02

**Exercise 1**

Provide the four assumptions of Poisson regression (as reviewed in lecture and in the book).

**Note**

Program facilitators keep track of the number of times their program's patients relapse within five years of initial treatment. Data on 100 patients yielded a mean number of 2.8 relapses per patient within the five years of initial treatment.

Use this prompt for exercises 2–5.

**Exercise 2**

Define the response.

**Exercise 3**

What are the possible values for the response?

**Exercise 4**

What does $\lambda$ represent?

**Exercise 5**

Would a zero-inflated model be considered here? If so, what would be a "true zero"?

**Exercise 6**

A Poisson regression model assumes that the mean and the variance ($\lambda$) are the same. Explain why or why not this assumption is realistic for observed data.

**Exercise 7**

Assume that our response variable is count data, however, we find that the empirical variances are larger than the empirical means. Provide two methods that provide a correction for this issue. (You can provide the name of the methods without describing them).

**Note**

A survey of 1,000 consumers asked respondents how many credit cards they use. Interest centers on the relationship between credit card use and income, in $10,000$. The estimated coefficient for income is 2.1.

Use this prompt for exercises 8-10.

**Exercise 8**

Identify the predictor and interpret the estimated coefficient in this context.

**Exercise 9**

Describe how the assumption of linearity can be assessed in this example.

**Exercise 10**

Suggest a valid way to assess the equal mean and variance assumption.

**Note**

Researchers are interested in the number of dates respondents arranged online and whether the rates differ by age group. Question which elicit responses similiar to this can be found in the Pew Survey concerning data online and relationships [Smith and Duggan, 2013]. Each survey respondent wa sasked how many dates they have arranged online in the past three months as well as the typical amount of time $t$ (in hours) they spend online weekly. Some rows of the data appear below.

| Age | Time Online | Number of Dates Arranged Online |
|---|---|---|
| 19 | 35 | 3 |
| 29 | 20 | 5 |
| 38 | 15 | 0 |
| 55 | 10 | 0 |

Use this prompt to answer exercises 11-13.

**Exercise 11**

Identify the response, predictor, and offset in this context. Does using an offset make sense?

**Exercise 12**

Write out a model for this data. As part of your model description, define the parameter, $\lambda$.

**Exercise 13**

Consider a zero-inflated Poisson model for this data. Describe what the 'true zeros' would be in this setting.