6: Generalized Linear Models

```
$ echo "Data Science Institute"
```

What happens if we are faced with a situation where the response Y is neither quantitative or qualitative?

Motivation

We have learned about linear and logistic regression which are generalized linear models (GLM). But exactly is GLM? Let's explore!

What is Generalized Linear Model?

According to Generalized Linear Modelling (GLM) is a flexible generalization of ordinary linear regression. Neither linear regression nor the classification approaches considered so far are applicable.

Examples of GLM

Here are three examples of GLM:

- Linear Regression
- Logistic Regression
- Poisson Regression

Poisson Regression

$$log(\lambda(X_1, \ldots, X_p) = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p$$

Note: Taking the log ensures that λ can only be non-negative.

This is equivalent to representing the mean λ as follows:

$$\lambda = \mathrm{E}(Y) = \lambda(X_1, \ldots, X_p) = e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p}$$

Exercise: Linear and Poisson Regression on Bikeshare data

Breakout Room

What are some advantages of Poisson Regression over Linear Regression?

Common Charactistics of GLM

- Use a set of predictors X_1, ..., X_p to predict a response Y
- Model the response Y as coming from a particular distribution

References

Chapter 4 of the ISLP book:

James, Gareth, et al. "Classification." An Introduction to Statistical Learning: with Applications in Python, Springer, 2023.