

Discrete Response Model

Lecture 5

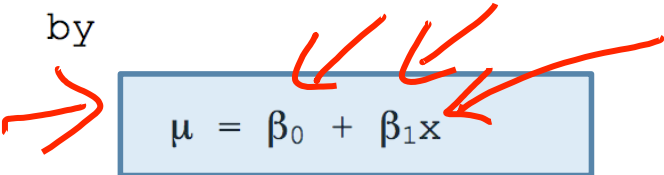
Models for Count Response, Discrete Response Model Evaluation, and Model Selection

datascience@berkeley


Poisson Regression Model: Model for Mean: Log-Link

Mean of the Poisson Distribution


Suppose the mean parameter of a Poisson distribution is now dependent on a function of explanatory variables. For example, suppose there is only one explanatory variable x . We could represent this dependence by


$$\mu = \beta_0 + \beta_1 x$$

Depending on the value of the parameters and x , we could obtain a negative value for μ which would not make sense for a count! Instead, we can use


$$\log(\mu) = \beta_0 + \beta_1 x$$

which alternatively can be written as


$$\mu = \exp(\beta_0 + \beta_1 x)$$

Now, μ is guaranteed to be greater than 0. This is referred to a Poisson regression model.

Mean as a Function of Explanatory Variables

When needed, we can emphasize that the mean changes as a function of the variable x for the i^{th} observation with

$$\mu_i = \exp(\beta_0 + \beta_1 x_i)$$

If there are p explanatory variables, we can write the model as

$$\mu = \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)$$

or

$$\log(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

Generalized Linear Model (GLM)

A Poisson regression model is a generalized linear model with the following components:

1. Random: Y has a Poisson distribution
2. Systematic: $\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$
3. Link: \log
 - A consequence of the log link function is that the explanatory variables affect the response mean in multiplicative way

Berkeley

SCHOOL OF
INFORMATION