

# Discrete Response Model

## Lecture 3

---

**[datascience@berkeley](mailto:datascience@berkeley)**

# Quadratic Term: An Introduction

# Incorporate Quadratic Terms in R

- Quadratic and higher-order polynomials are one way to capture the nonlinear relationship between an explanatory variable and  $\text{logit}(\pi)$ .
- To include this type of transformation in a formula argument, we can use the carat symbol ^ with the degree of the polynomial.
- However, as we saw earlier in this section, the carat symbol is used to denote the order of the interaction between explanatory variables. Thus,

```
formula = y ~ x1 + x1^2
```

would NOT be interpreted as  $\text{logit}(\pi) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$ .

- R interprets the  $x1^2$  part as "all two-way interactions involving  $x1$ ." Because only one explanatory variable is given in  $x1^2$ , R interprets this as simply  $x1$ . Also, because  $x1$  was already given in the formula argument, the variable is not duplicated, so  $\text{logit}(\pi) = \beta_0 + \beta_1 x_1$  is estimated instead!

# Incorporate Quadratic Terms in R

- In order to obtain a  $x_1^2$  terms, we need to use the `I()` function with `x1^2`.
- The `I()` function instructs R to interpret arguments as it normally would.
- Thus, `formula = y ~ x1 + I(x1^2)` would be interpreted as

$$\text{logit}(\pi) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$$

# Berkeley

SCHOOL OF  
INFORMATION