



DS-UA 112

Introduction to Data Science

Week 13: Lecture 2

Logistic Regression





How can we modify linear regression to predict qualitative variables?

DS-UA 112

Introduction to Data Science

Week 13: Lecture 1

Logistic Regression

Adapted from Nolan, Speed, Gonzalez, Lau



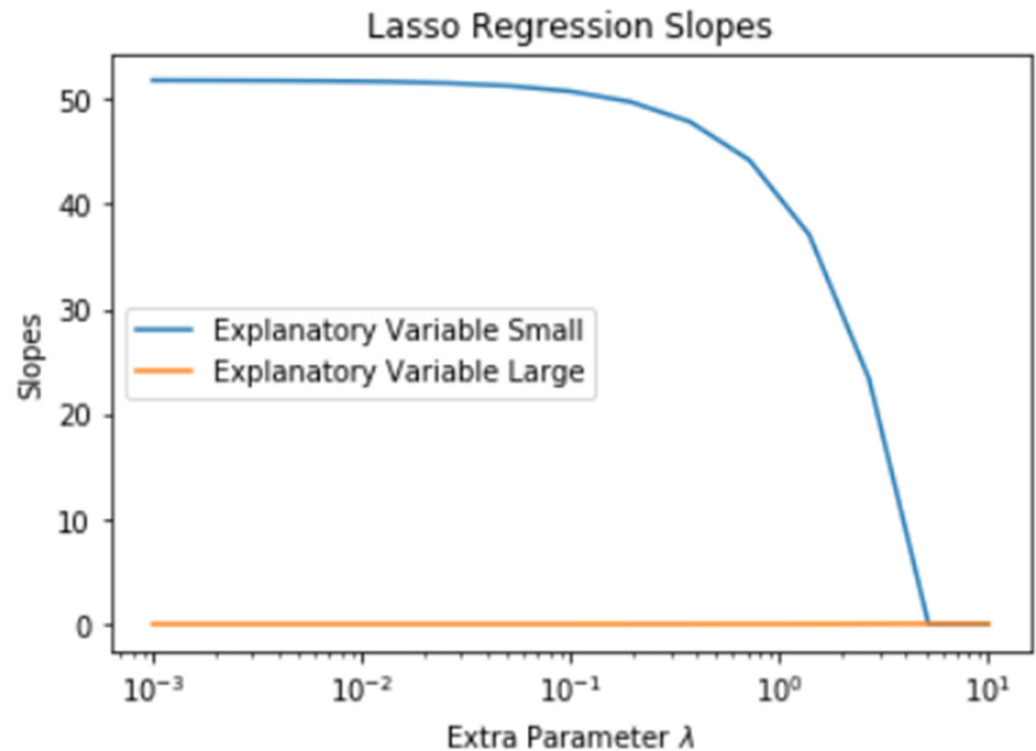
Announcements

- ▶ Please check Week 13 agenda on NYU Classes
 - ▶ Lab 12
 - ▶ Due on Friday April 24 at 11:59PM EST
 - ▶ Homework 5
 - ▶ Due on Friday May 1 at 11:59PM EST



Review

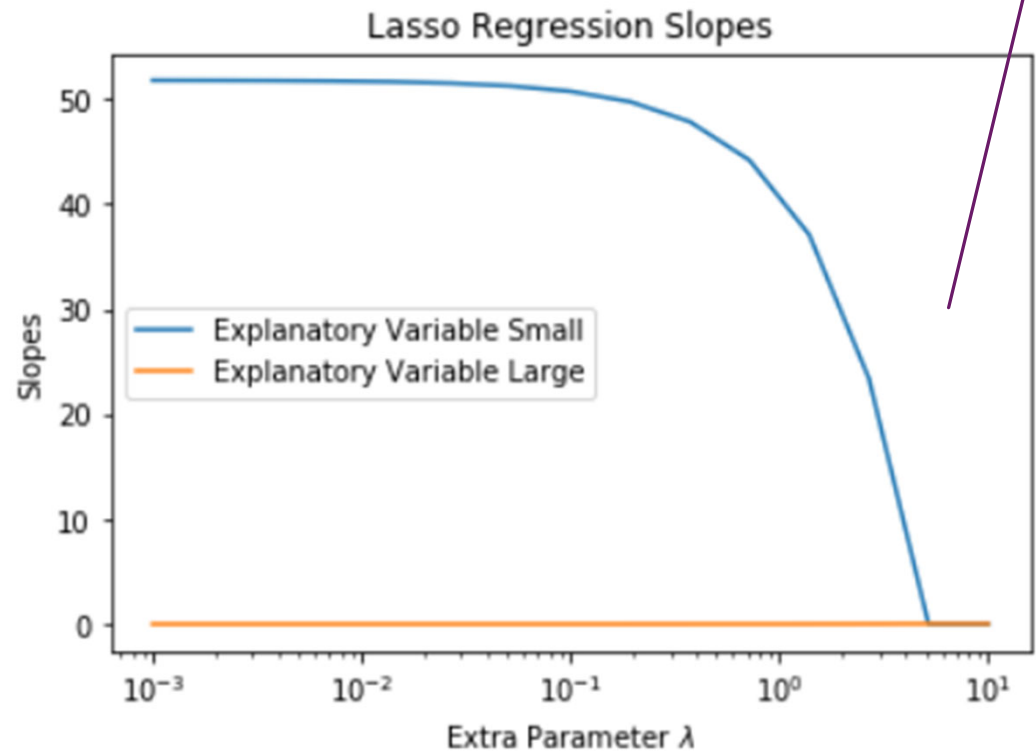
- ▶ Remember that we have pairs of features and parameters in models.
 - ▶ If the features are large then the parameters tend to be small.
 - ▶ If the features are small then the parameters tend to be large.
- ▶ If the features have different scales then regularization tends to shrink parameters for small features more than parameters for large features



Review

Regularization requires transforming to standard units

- ▶ Remember that we have pairs of features and parameters in models.
 - ▶ If the features are large then the parameters tend to be small.
 - ▶ If the features are small then the parameters tend to be large.
- ▶ If the features have different scales then regularization tends to shrink parameters for small features more than parameters for large features



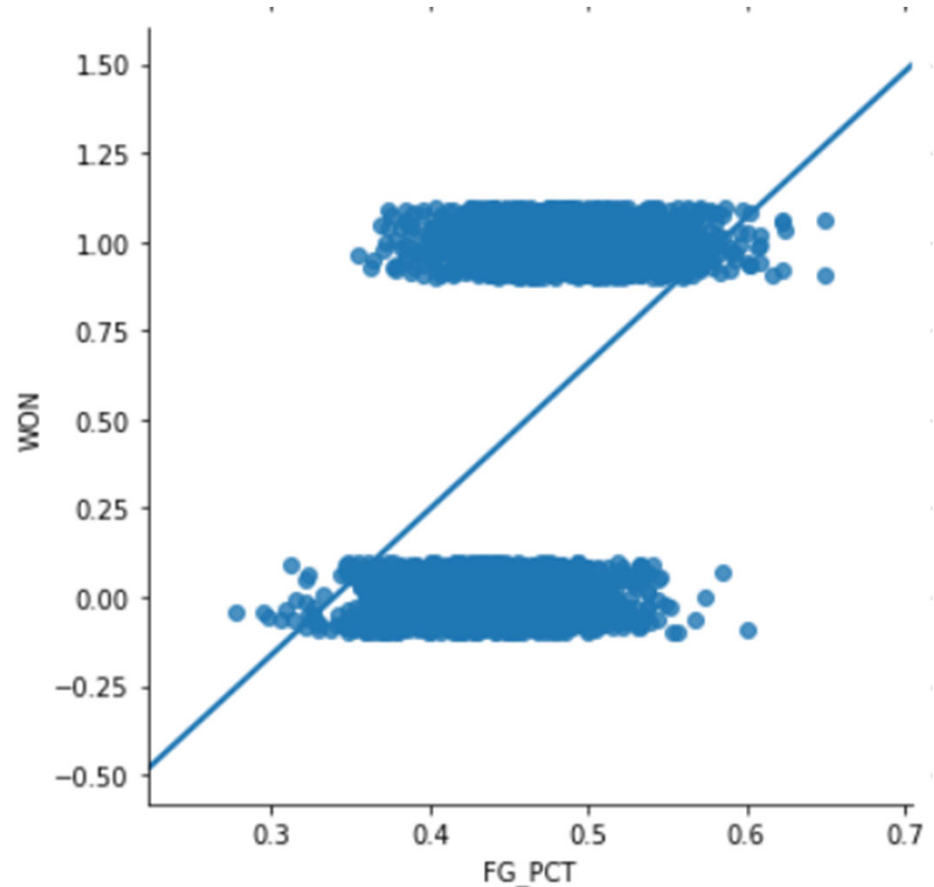
Agenda

- ▶ Logistic Regression
 - ▶ Sigmoid Function
 - ▶ Logistic Loss

applied
algorithm don't
interest understanding deep
statistics learning field
program learn lot idea
model fun set skill job
expect work data world
gain good science tool large
project ds hope
knowledge real python
basic application method
class making practical analyze
experience library help
create expand actual

Qualitative Data

- ▶ Suppose we want to predict a variable consisting of qualitative data. If we have two categories, then we can encode the categories as 1 and 0.
- ▶ Linear regression cannot determine predictions because lines cannot fit the pattern in the data

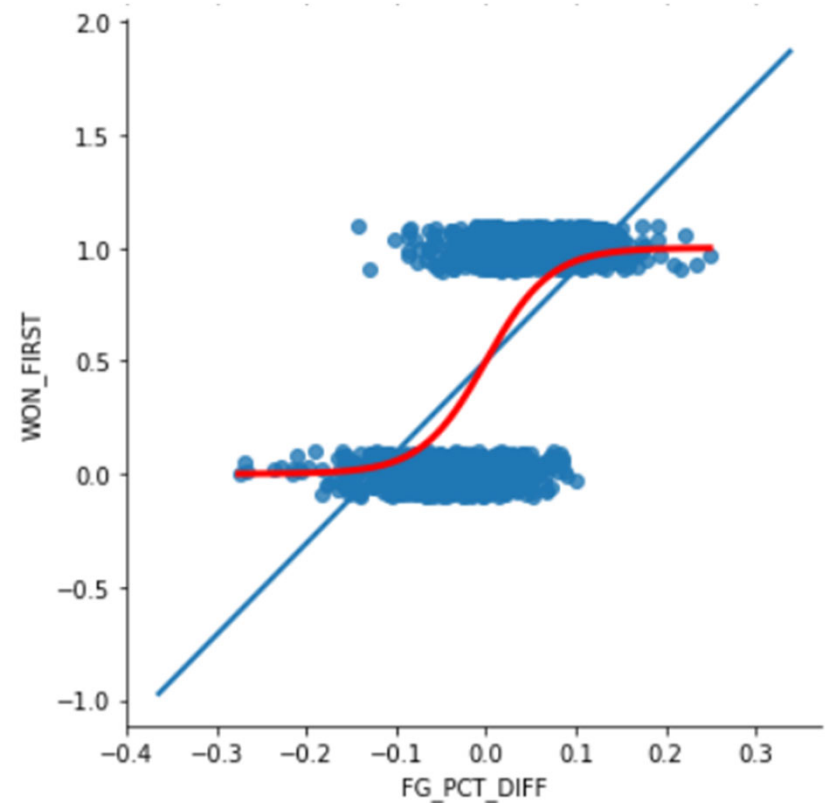


Sigmoid Function

- Instead we can use the **sigmoid function** which makes a shape like the letter S

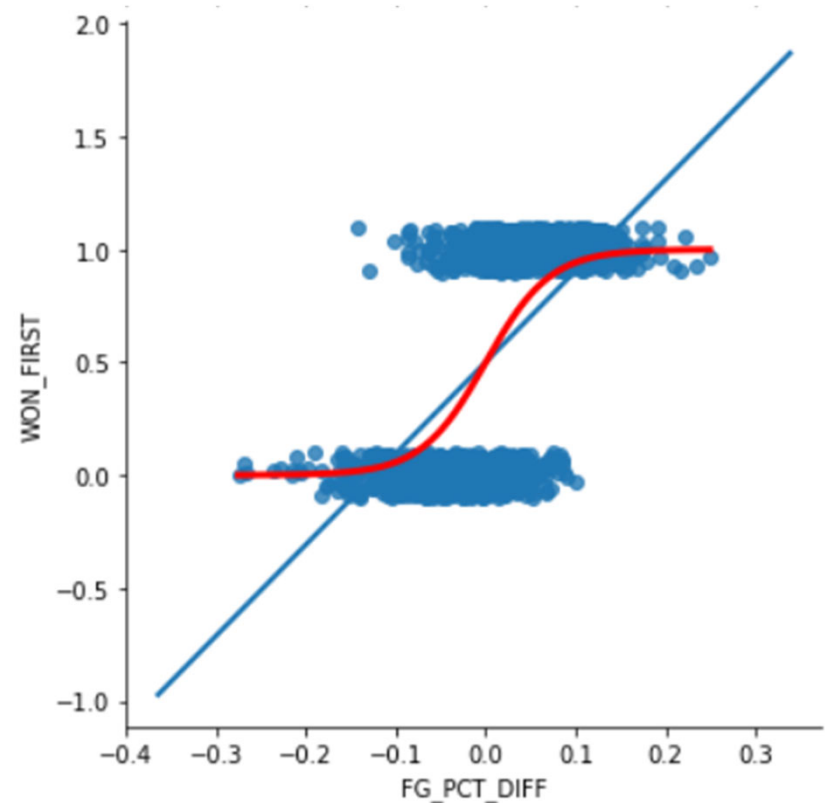
$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

- If we replace t with $w * t$ for some number w , then can adjust the shape to match the data



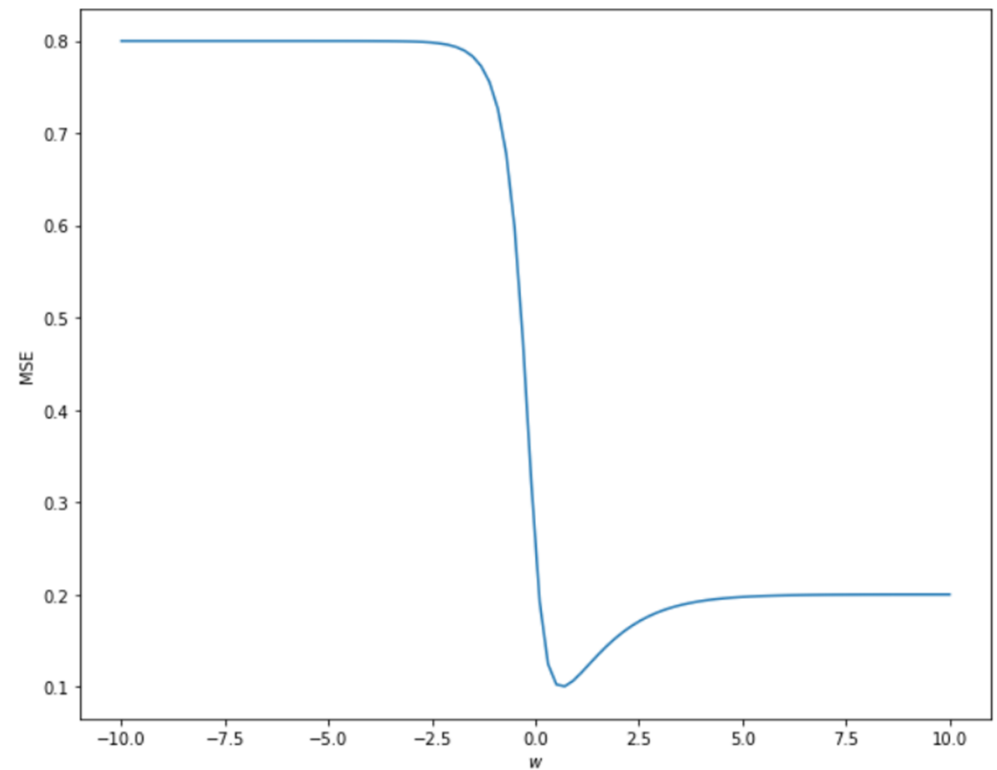
Logistic Regression

- ▶ In **logistic regression** we predict numbers between 0 and 1 to match the categories
- ▶ We could interpret the numbers between 0 and 1 like probabilities.
- ▶ Conditional on the value of the explanatory variable we predict the probability that the response variable is category 0 or category 1.



Logistic Loss

- ▶ While we can use the same approaches from linear regression such as gradient descent to determine the parameters in logistic regression we should use a different loss function.
- ▶ Instead of the square loss we should take the **logistic loss**. Unlike the square loss, the logistic loss does not generate flat regions that prevent gradient descent from finding the minimum.



Summary

- ▶ Logistic Regression
 - ▶ Sigmoid Function
 - ▶ Logistic Loss

Goals

- ▶ Understand the inability of linear regression to model qualitative variables
- ▶ Make a connection between sigmoid function and odds ratio
- ▶ Use logistic loss instead of square loss to compare observations and predictions

Questions

- ▶ Questions on Piazza?
 - ▶ Please provide your feedback along with questions
- ▶ Question for You!
 - ▶ Should we use knowledge about a population to determine models?

NEW SCIENTIST

Specialist Knowledge Is Useless and Unhelpful

When data prediction is a game, the experts lose out.

By PETER ALDHOUS

DEC 08, 2012

PA: That sounds very different from the traditional approach to building predictive models. How have experts reacted?

JH: The messages are uncomfortable for a lot of people. It's controversial because we're telling them: "Your decades of specialist knowledge are not only useless, they're actually unhelpful; your sophisticated techniques are worse than generic methods." It's difficult for people who are used to that old type of science. They spend so much time discussing whether an idea makes sense. They check the visualizations and noodle over it. That is all actively unhelpful.

