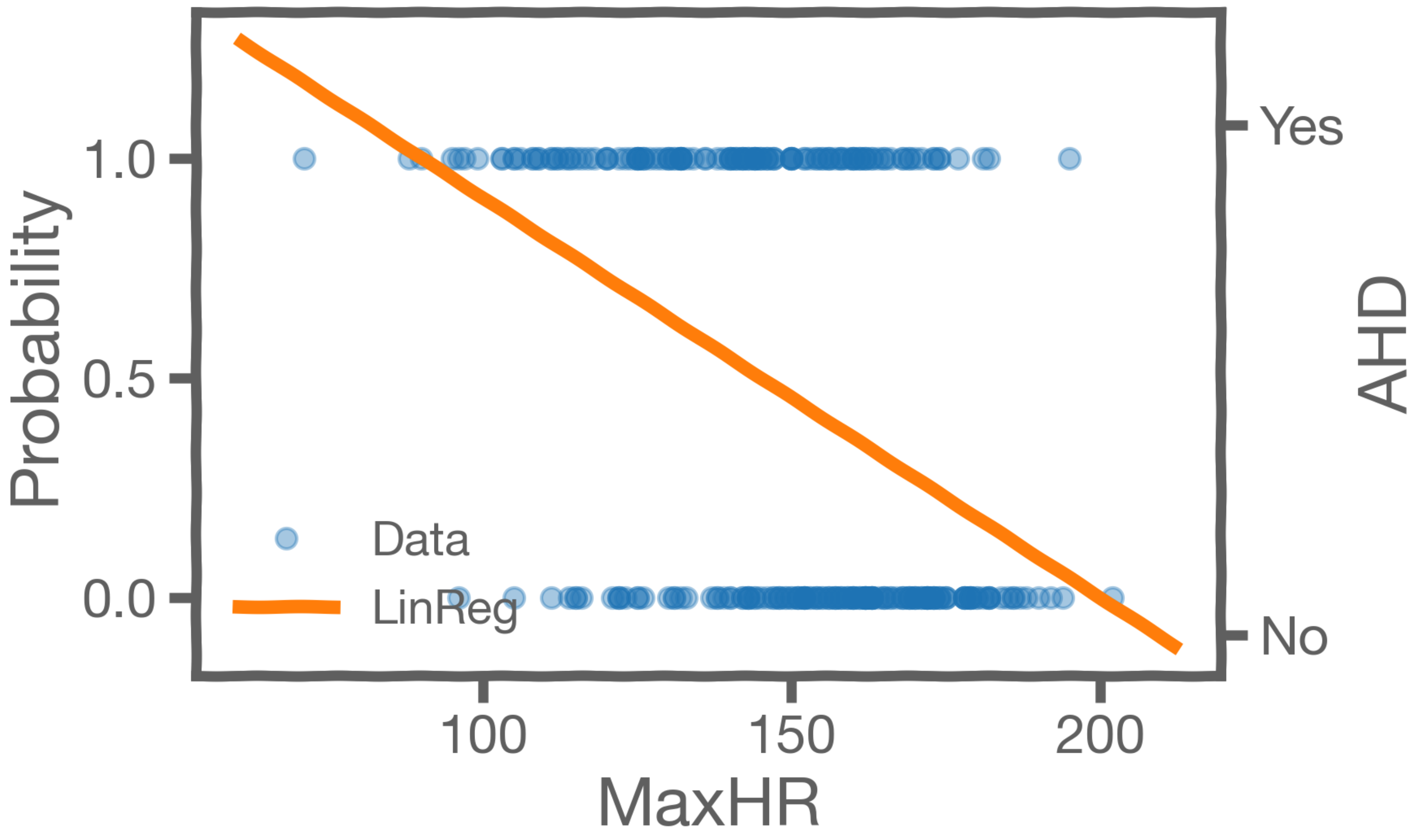
Logistic Regression

Bookmark this page

Could a linear function perform classification?

We can use linear functions to do regression. But suppose we want to take a patient's maximum heart rate and predict a discrete variable: whether or not they can be diagnosed with heart disease. Linear regression could learn a function that "fits" the data in some sense - but the shape of a linear model doesn't fit our discrete dataset well. For example, for data further to the left or right it would output a value beyond the range (0,1), which cannot be interpreted as a probability.



Therefore we would like to change the shape of our model to look more discrete.

Can we create a continuous function that is a close approximation to a discrete function?

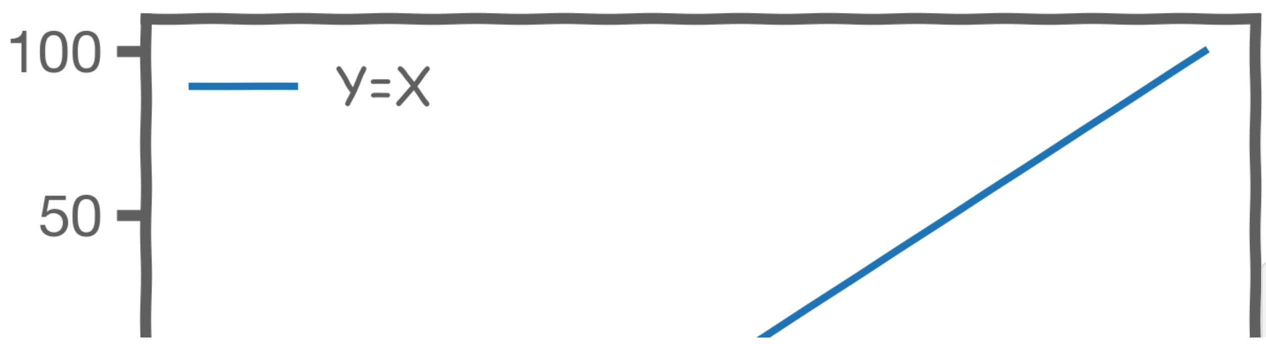
$$\text{discrete}\,(x) = \left\{ \begin{smallmatrix} 0 & \text{ifj } x<0 \\ 1 & \text{otherwise} \end{smallmatrix} \right\}$$

What is a logistic function?

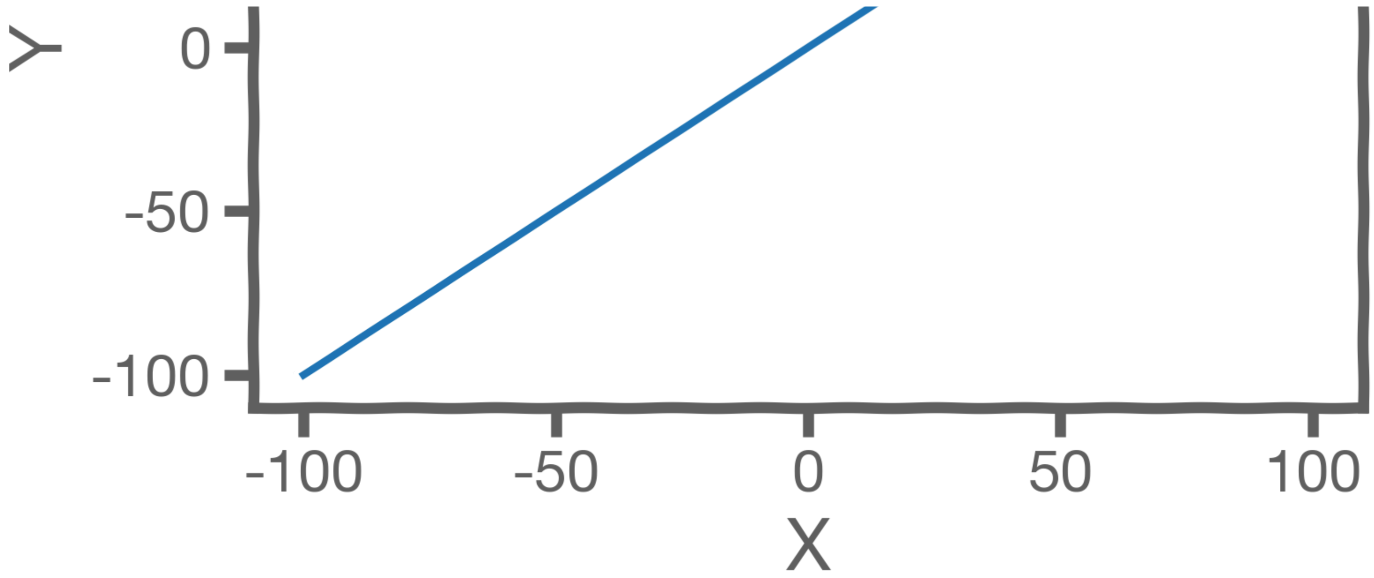
> **Logistic Functions**
>
> The formula for the kind of logistic function above is given by the sigmoid function:
>
> $$f\,(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

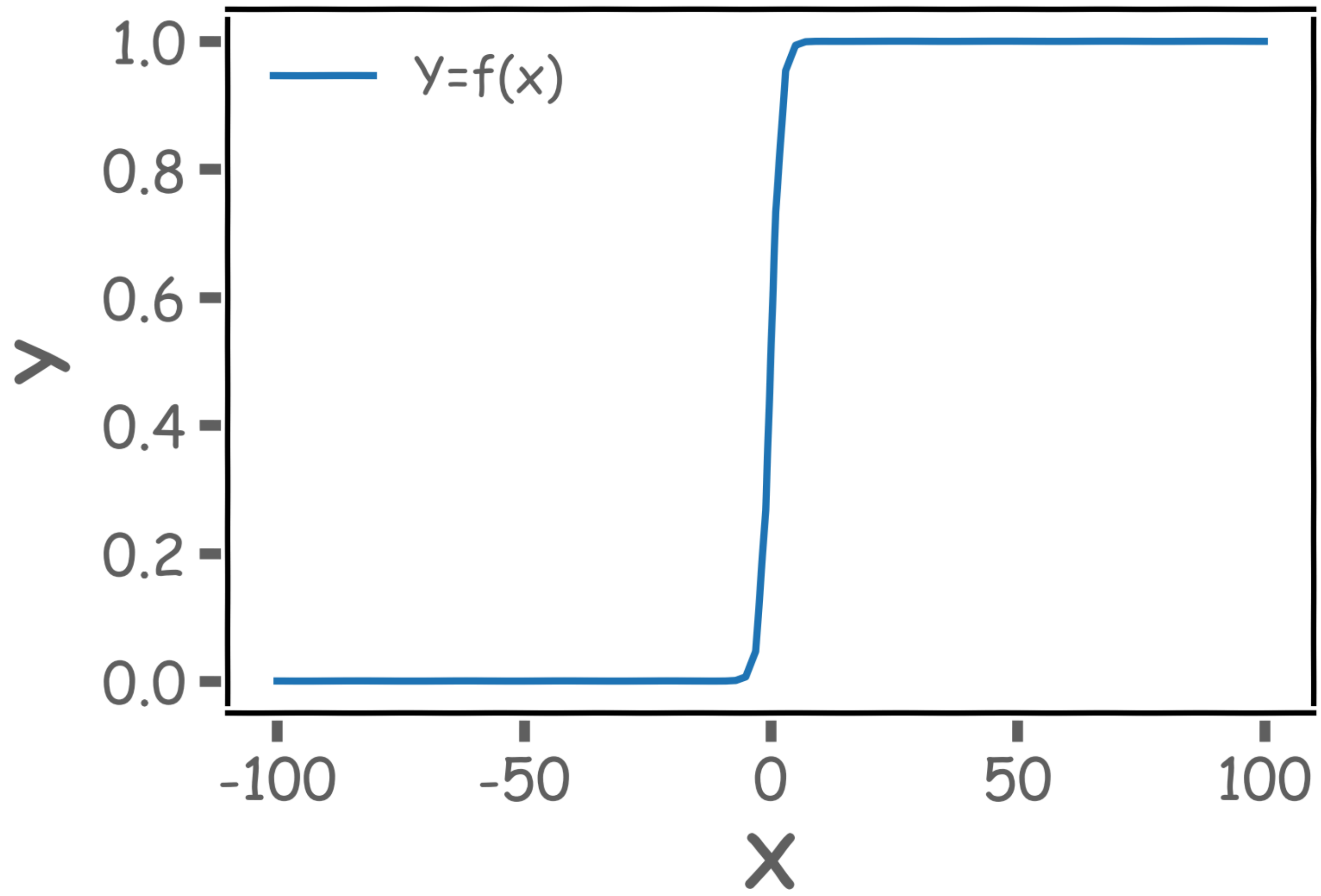
And by learning the optimal values of the coefficients β_0 and β_1, we can find the function of a curve to label binary data.

Here is what it looks like to transform a linear function with the sigmoid function:



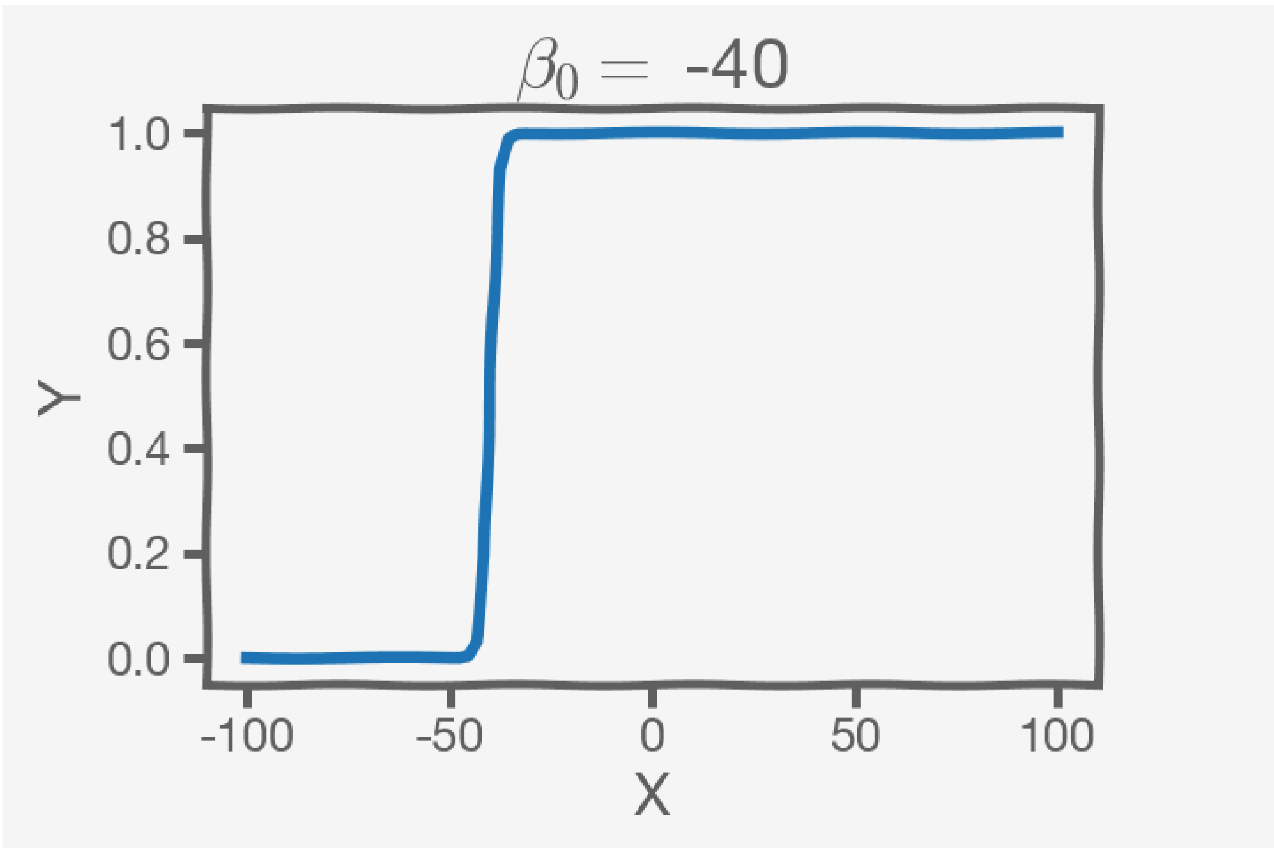
Calculator



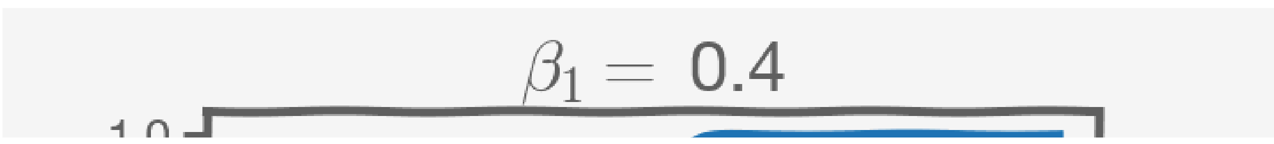


Adjusting our Beta values adjusts the shape of the logistic function

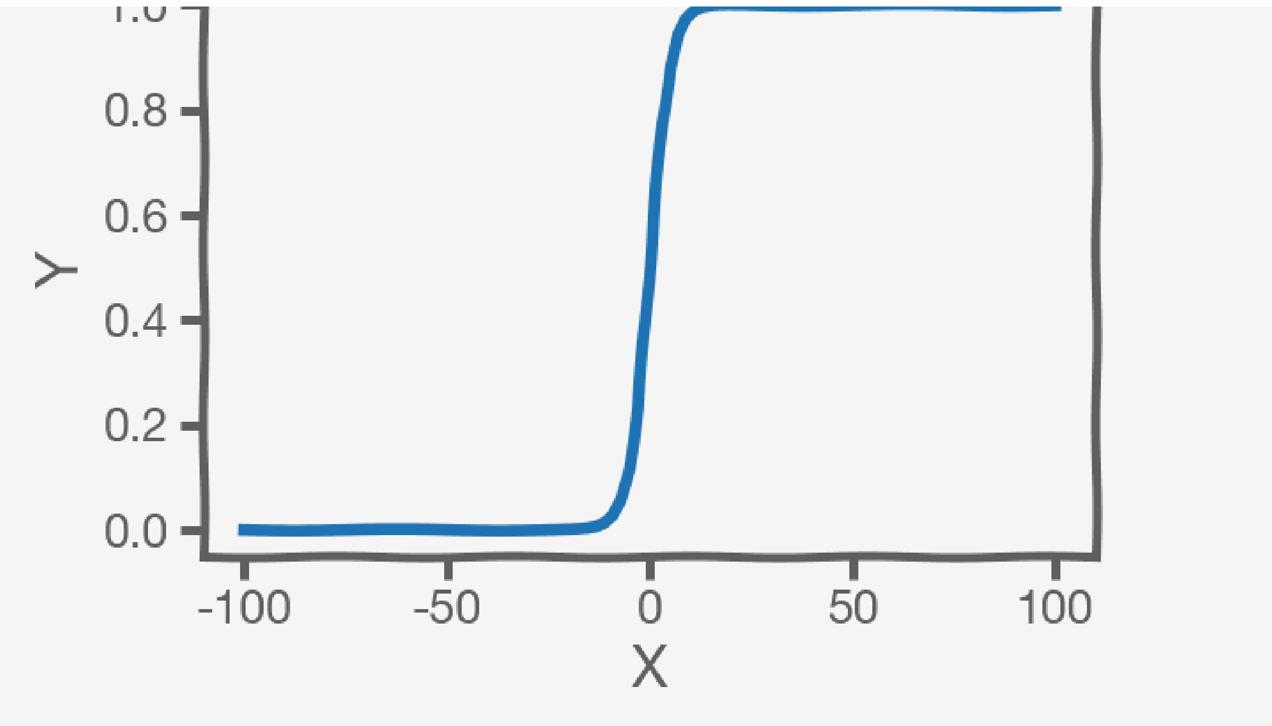
Changing β_0



Changing β_1

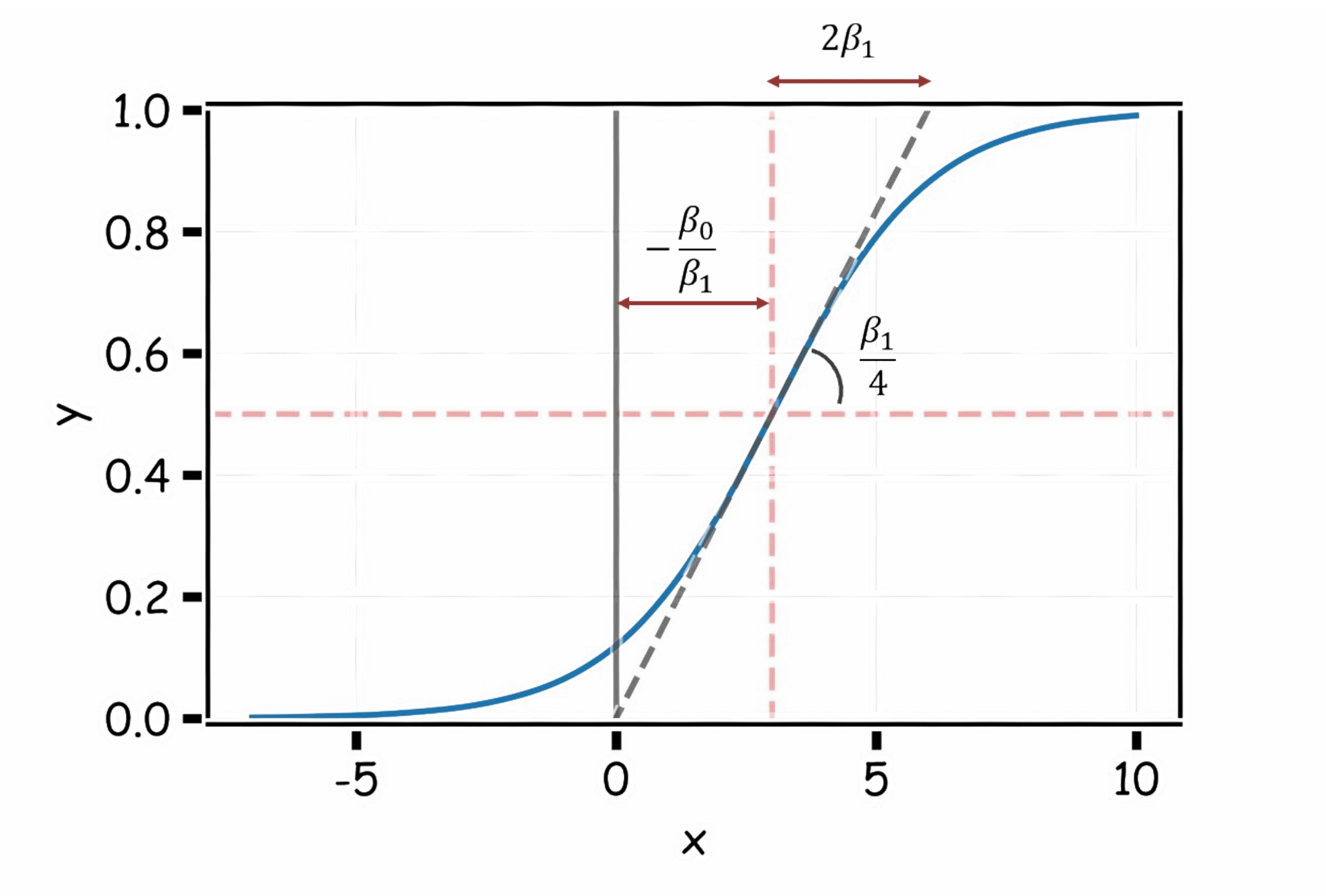


Calculator



β_0 shifts the logistic curve right or left (scaled by β_1, as seen in the arrow towards the middle of the diagram below). β_1 controls how steep the logistic curve is. The distance from the middle to the top (AKA the distance of the curve from outputting a probability of 0.5 to outputting of almost 1.0) is $2|\beta_1$. So when the absolute value of β_1 increases and our curve gets steeper, almost all our output probabilities move towards 0 and 1. But when the absolute value of β_1 is smaller, that means there are many data points on which are model loses confidence and outputs more probabilities closer to 0.5.

<



If β_1 is positive, then the predicted probability increases from low to high as we increase our input. If β_1 is negative, then the predicted probability decreases from high to low as we increase our input.

Interpreting Beta values as odds

Since we are interpreting the output of our logistic function (the sigmoid function) as a probability

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_i)}}$$

We can rearrange the equation in terms of a **log-odds ratio**. The following equation is equivalent to the one above:

$$\log\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 x_i$$

The ratio of probabilities

Calculator

$$\frac{p_i}{1-p_i}$$

is called the **odds** term, which rises and falls along with the probabilities themselves, just on a different scale. The probability is between 0 and 1, but the odds are between 0 and infinity. (For example, when the probability is 0.25 the odds are 1/3; when the probability is 0.5 the odds are 1; when the probability is 0.75 the odds are 3; when the probability is 0.9 the odds are 9; when the probability is 0.995 the odds are 199, etc.) The log of this ratio is called the **log-odds**.

Now we can interpret how β_0 and β_1 change how the model responds to our input.

β_0 is the **log-odds** our model outputs when $x = 0$.

Equivalently

$$\frac{e^{\beta_0}}{1+e^{\beta_0}}$$

is the **probability** our model outputs when $x = 0$.

One unit of change in x_i corresponds to β_1 units of change in the log-odds our model outputs. That is why is it sometimes useful to consider log-odds instead of probabilities, because modeling the probabilities is equivalent to modeling the log-odds.

Other methods?

You might have noticed in this course that there's never just one way to do things. In the video below, Max talks about the ROC curve classification method, and how he used it in a project involving US politics.

Video

Video

Descargar archivo de video

Transcripciones

Descargar SubRip (.srt) file
Descargar Text (.txt) file

Calculator

Discussion Board (External resource)

Haga clic en Aceptar para que su nombre de usuario y dirección de correo electrónico se envíen a una aplicación de terceros.

Aceptar

< Previous Next >

© All Rights Reserved

© 2024 edX LLC. All rights reserved.
深圳市恒宇博科技有限公司 粤ICP备17044299号-2

Calculator