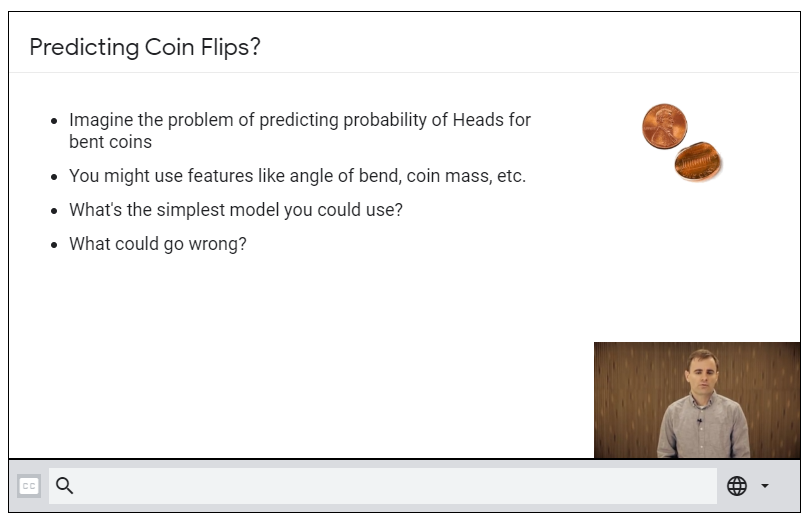
Logistic Regression

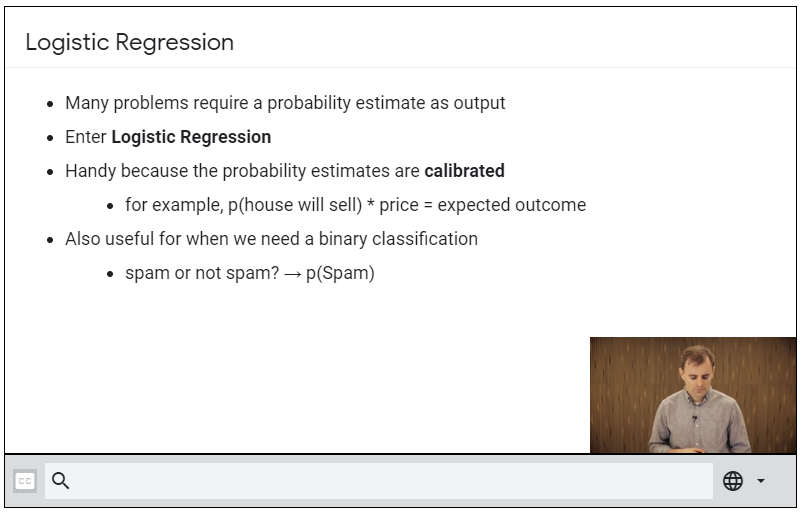
Instead of predicting exactly 0 or 1, **logistic regression** generates a probability—a value between 0 and 1, exclusive. For example, consider a logistic regression model for spam detection. If the model infers a value of 0.932 on a particular email message, it implies a 93.2% probability that the email message is spam. More precisely, it means that in the limit of infinite training examples, the set of examples for which the model predicts 0.932 will actually be spam 93.2% of the time and the remaining 6.8% will not.

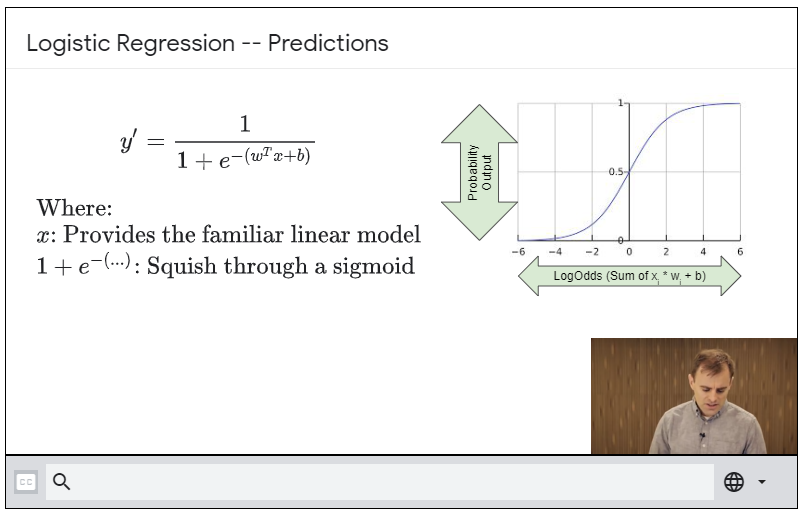


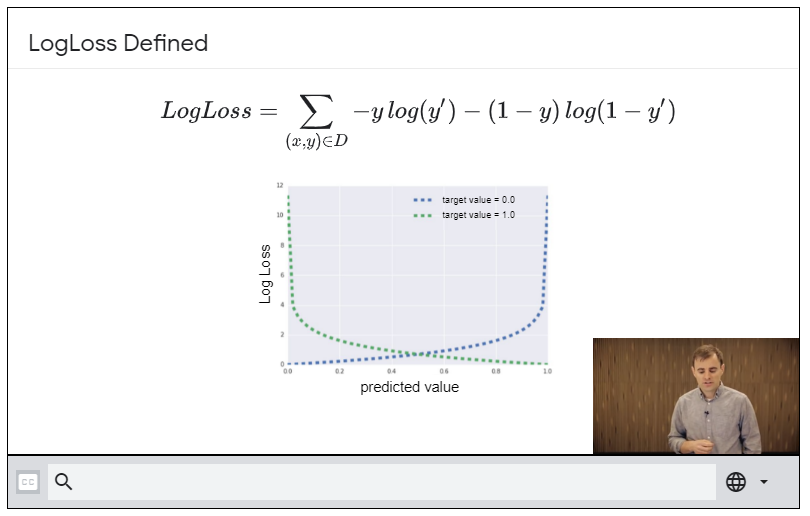
But problem using linear regression for the predicting probability of getting a HEAD in coin toss is that:

Consider a new coin with more mass, different angle of bend etc. Then it’s possible that our model could end predicting a value that is not in between 0 and 1.

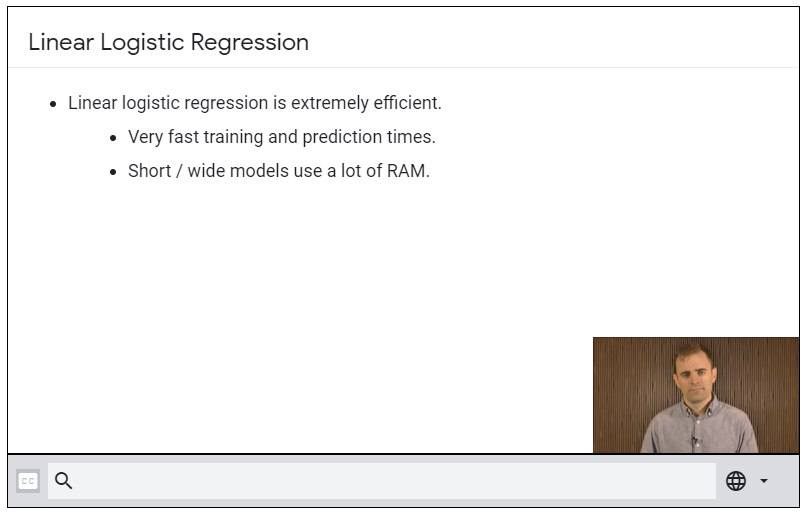
Since probability always needs to be between 0 and 1, this can’t be right.











If we need non-linearity, we can always use feature-crosses (remember using x1.x2, x12 and x22 to form circle-kind of predicting models?)