*1*) How would you make a prediction using a *Logistic Regression* model?

*2)*When *can Logistic Regression* be used?

*3*)Why is Logistic Regression called *Regression* and not *Classification*?

*4*)How is a Logistic *Regression* model trained?

*5*)Provide a mathematical intuition for Logistic Regression?

*6*)What can you infer from each of the hand drawn decision boundaries of *Logistic Regression* below?

*7*)What is the difference between *Linear Regression* and *Logistic Regression*?

*8*) Why can't a *Linear Regression* be used instead of *Logistic Regression*?

*9)*Why don’t we use *Mean Squared Error* as a cost function in Logistic Regression?

*10*)Why is *Logistic Regression* considered a *Linear Model*?

**LOGISTIC REGRESSION**

## *Q1*: How would you make a prediction using a *Logistic Regression* model?

Answer

In Logistic regression models, we are modeling the *probability* that an input (X) belongs to the default class (Y=1), that is to say:

*P*(*X*)=*P*(*Y*=1∣*X*)

where the P(X) values are given by the *logistic function*,

​

The β0 and β1 values are estimated during the training stage using *maximum-likelihood* estimation or *gradient descent*. Once we have it, we can make predictions by simply putting numbers into the *logistic regression equation* and calculating a result.

For example, let's consider that we have a model that can predict whether a person is male or female based on their height, such as if P(X) ≥ 0.5 the person is male, and if P(X) < 0.5 then is female.

During the training stage we obtain β0 = -100 and β1 = 0.6, and we want to evaluate what's the probability that a person with a height of 150cm is male, so with that intention we compute:

Given that logistic regression solves a *classification* task, we can directly use this value to predict that the person is a female.

## *Q2*: When *Logistic Regression* can be used?

Answer

Logistic regression can be used in *classification* problems where the output or dependent variable is *categorical* or *binary*. However, in order to implement logistic regression correctly, the dataset must also satisfy the following properties:

1. There should not be a high correlation between the independent variables. In other words, the predictor variables should be independent of each other.
2. There should be a linear relationship between the logit of the outcome and each predictor variable. The logit function is given as logit(p) = log(p/(1-p)), where p is the probability of the outcome.
3. The sample size must be large. How large depends on the number of independent variables of the model.

When all the requirements above are satisfied, logistic regression can be used.

## *Q3*: Why is Logistic Regression called *Regression* and not *Classification*?

Answer

Although the task we are targeting in logistic regression is a classification, *logistic regression does not actually individually classify things for you*: it just gives you probabilities (or log odds ratios in the logit form).

The only way logistic regression can actually classify stuff is if you apply a rule to the probability output. For example, you may round probabilities greater than or equal to 50% to 1, and probabilities less than 50% to 0, and that’s your classification.

## *Q4*: How a *Logistic Regression* model is trained?

Answer

The logistic model is trained through the logistic function, defined as:

p(y)= 1/ (1+ )

​

where x is the input data, w is the weight vector, y is the output label, and P(y) is the probability that the output label belongs to one class. If for some input we got P(y) > 0.5, then the predicted output is 1, and otherwise would be 0.

The training is based in estimate the w vector. For this, in each training instance we use *Stochastic Gradient Descent* to calculate a prediction using some initial values of the coefficients, and then calculate new coefficient values based on the error in the previous prediction. The process is repeated for a fixed number of iterations or until the model is accurate enough or cannot be made any more accurate.

## *Q5*: Provide a mathematical intuition for Logistic Regression?

Answer

Logistic regression can be seen as a transformation from linear regression to logistic regression using the logistic function, also known as the sigmoid function or S(x):

​

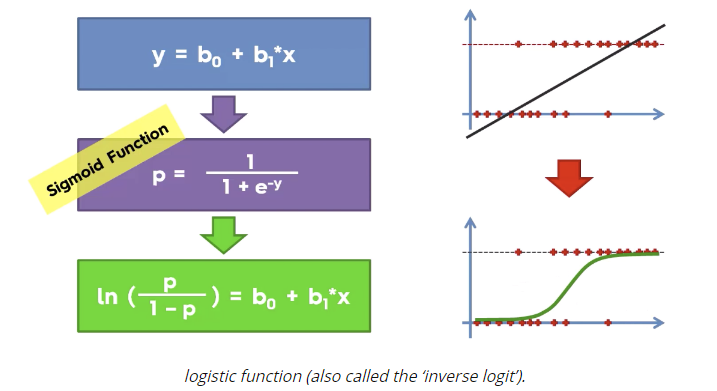
Given the linear model:

If we apply the sigmoid function to the above equation it results:

where p is the probability and it takes values between 0 and 1. If we now apply the logit function to p, it results:

The equation above represents the logistic regression. It fits a logistic curve to set of data where the dependent variable can only take the values 0 and 1.

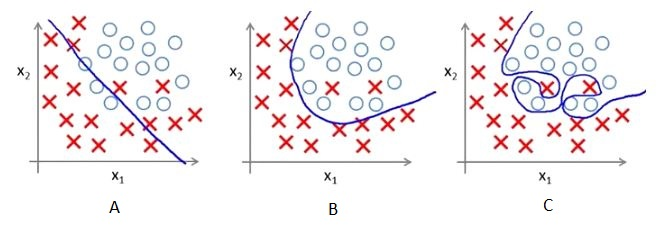
The previous transformation can be illustrated in the following figure:



## *Q6*: What can you infer from each of the hand drawn decision boundary of *Logistic Regression* below?

Problem

Also, what should we do to fix the problem of each decision boundary?



Answer

What can we infer:

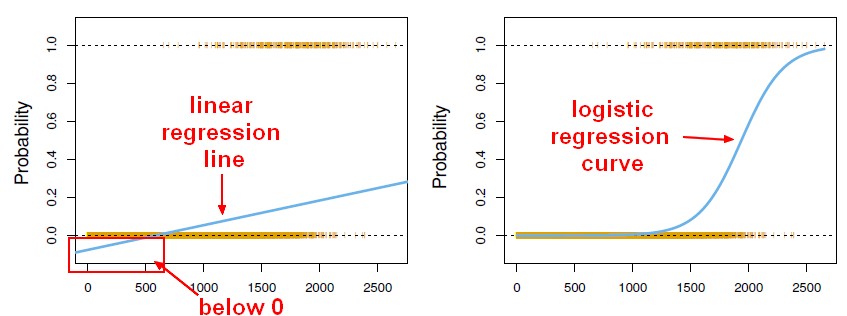
* A: the model *underfits* the data. It will give us the maximum error compared to other two models.
* B: *best-fitting* model.
* C: the model *overfits* the data. It performs exceptionally well on training data but performs considerably worse on test data.

What can we do to fix the problem:

* A: *increase* the complexity of the model or *increase* the number of independent variables.
* B: best performing model, so we don't need to tweak anything.
* C: add *regularization* method to the model.

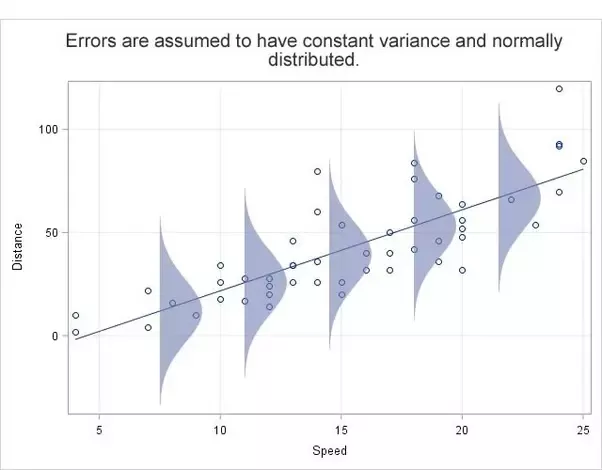
## *Q7*: What is the difference between *Linear Regression* and *Logistic Regression*?

Answer

* Linear regression output as probabilities In linear regression, the outcome (dependent variable) is continuous. It can have any one of an infinite number of possible values. In logistic regression, the outcome (dependent variable) has only a limited number of possible values.  
  
* Outcome In linear regression, the outcome (dependent variable) is continuous. It can have any one of an infinite number of possible values. In logistic regression, the outcome (dependent variable) has only a limited number of possible values.
* The dependent variable Logistic regression is used when the response variable is categorical in nature. For instance, yes/no, true/false, red/green/blue, 1st/2nd/3rd/4th, etc. Linear regression is used when your response variable is continuous. For instance, weight, height, number of hours, etc.
* Equation Linear regression gives an equation which is of the form Y = mX + C, means equation with degree 1. However, logistic regression gives an equation which is of the form Y = eX + e-X
* Coefficient interpretation In linear regression, the coefficient interpretation of independent variables are quite straightforward (i.e. holding all other variables constant, with a unit increase in this variable, the dependent variable is expected to increase/decrease by xxx). However, in logistic regression, depends on the family (binomial, Poisson, etc.) and link (log, logit, inverse-log, etc.) you use, the interpretation is different.
* Error minimization technique Linear regression uses *ordinary least squares* method to minimise the errors and arrive at a best possible fit, while logistic regression uses *maximum likelihood* method to arrive at the solution. Linear regression is usually solved by minimizing the least squares error of the model to the data, therefore large errors are penalized quadratically. Logistic regression is just the opposite. Using the logistic loss function causes large errors to be penalized to an asymptotically constant. Consider linear regression on categorical {0, 1} outcomes to see why this is a problem. If your model predicts the outcome is 38, when the truth is 1, you've lost nothing. Linear regression would try to reduce that 38, logistic wouldn't (as much)[2](https://stats.stackexchange.com/a/29340).

## *Q8*: Why can't a *Linear Regression* be used instead of *Logistic Regression*?

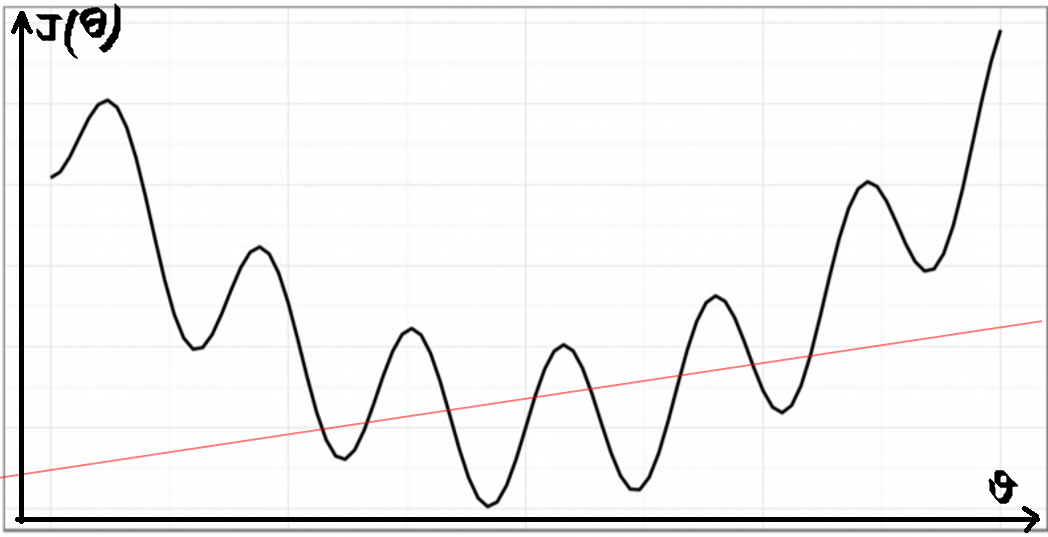
Answer

* It is required for the independent and dependent variables to be linear for linear regression models, but the independent and dependent variables are not required to have a linear relationship in logistic functions.
* The Linear Regression models assume that the error terms are *normally distributed* (bell-shaped graph) whereas there are *no error terms* in Logistic Regression because it is assumed to follow a *Bernoulli distribution*.
* Linear regression has a *continuous output*. Logistic regression does *not* have a continuous output, rather the output is a probability between 0 and 1. A linear regression may have an output that can go beyond 0 and 1
* 

## *Q9*: Why don’t we use *Mean Squared Error* as a cost function in Logistic Regression?

Answer

* In Linear Regression, we used the Squared Error mechanism.
* For Logistic Regression, such a cost function produces a non-convex space with many local minimums, in which it would be very difficult to minimize the cost value and find the global minimum.



## *Q10*: Why is *Logistic Regression* considered a *Linear Model*?

Answer

A model is considered linear if the transformation of features that is used to calculate the prediction is a linear combination of the features. Although Logistic Regression uses Sigmoid function which is a nonlinear function, the model is a generalized linear model because the outcome always depends on the sum of the inputs and parameters.

i.e the logit of the estimated probability response is a linear function of the predictors parameters.

