**Evaluating a Hypothesis**

Once we have done some trouble shooting for errors in our predictions by:

* Getting more training examples
* Trying smaller sets of features
* Trying additional features
* Trying polynomial features
* Increasing or decreasing λ

We can move on to evaluate our new hypothesis.

A hypothesis may have a low error for the training examples but still be inaccurate (because of overfitting). Thus, to evaluate a hypothesis, given a dataset of training examples, we can split up the data into two sets: a **training set** and a **test set**. Typically, the training set consists of 70 % of your data and the test set is the remaining 30 %.

The new procedure using these two sets is then:

1. Learn Θ and minimize Jtrain(Θ) using the training set
2. Compute the test set error Jtest(Θ)

**The test set error**

1. For linear regression: 
2. For classification ~ Misclassification error (aka 0/1 misclassification error):



This gives us a binary 0 or 1 error result based on a misclassification. The average test error for the test set is:



This gives us the proportion of the test data that was misclassified.