**Lasso Regression:**

Lasso regression is a type of linear regressionthat uses shrinkage. Shrinkage is where data values are shrunk towards a central point, like the mean. The lasso procedure encourages simple, sparse models . This particular type of regression is well-suited for models showing high levels of multi collinearity or when you want to automate certain parts of model selection, like variable selection/parameter elimination.

The acronym “LASSO” stands for Least Absolute Shrinkage and Selection Operator.

**L1 Regularization:**

Lasso regression performs L1 regularization, which adds a penalty equal to the absolute value of the magnitude of coefficients. This type of regularization can result in sparse models with few coefficients, some coefficients can become zero and eliminated from the model. Larger penalties result in coefficient values closer to zero, which is the ideal for producing simpler models. On the other hand, L2 regularization doesn’tresult in elimination of coefficients or sparse models. This makes the Lasso far easier to interpret than the Ridge.

**Ridge Regression:**

Ridge Regression is a technique for analysing multiple regression data that suffer from multi collinearity. When multi collinearity occurs, least squares estimates are unbiased, but their variances are large so they may be far from the true value. By adding a degree of bias to the regression estimates, ridge regression reduces the standard errors. It is hoped that the net effect will be to give estimates that are more reliable. Another biased regression technique, principal components regression, is also available in NCSS. Ridge regression is the more popular of the two methods.

* It shrinks the parameters, therefore it is mostly used to prevent multi collinearity.
* It reduces the model complexity by coefficient shrinkage.
* It uses L2 regularization technique.

**Elastic Net Regression:**

The method linearly combines the L1 and L2 penalties of the LASSO and Ridge Regression. Including the Elastic Net, these methods are especially powerful when applied to very large data where the number of variables might be in the thousands or even millions.

In mathematics, **sparse** and **dense** often refer to the number of zero vs. non-zero elements in an array (e.g. vector or matrix). A sparse array is one that contains mostly zeros and few non-zero entries, and a dense array contains mostly non-zeros. LASSO and Ridge encourage sparse and dense model, respectively, but since it never be that clear how the true model looks like, it’s typical to apply both methods and determine the best model.