Dan Bilsker

Project Summary

This multi-part project uses machine learning to construct a model to predict maximum drawdown(MDD) in stock prices values.

In the first part, we use a real time series from the financial market as data and divide the close values into “epochs,” which are periods of equal length within which we compute our MDD. We can then plot a histogram and probability distributions of the MDD over a range of epochs. The 90MDD is taken as a data point of analytic interest – 90% of MDD values are to the left of the 90MDD on the density distribution.

In the second part, we use Python to create a synthetic time series. The synthetic data is a Wiener process that models the original financial data. We take 2048 time series and take the MDD over different epochs.

Finally, we use the synthetic time series as training and validation data for a machine learning model in order to predict the MDD. To do this we begin by standardizing the data, then we apply a linear regression and polynomial regression model. To avoid overfitting, we use cross-k validation. We then test with the real world data and evaluate the model.

We can rank validation results for each polynomial degree, graph the coefficients of our model, and plot the relation between standard deviation and 90MDD. The model has a high rate of accuracy and indicates a strong relationship between standard deviation and maximum drawdown.