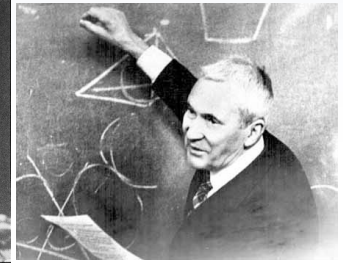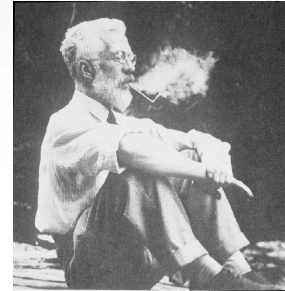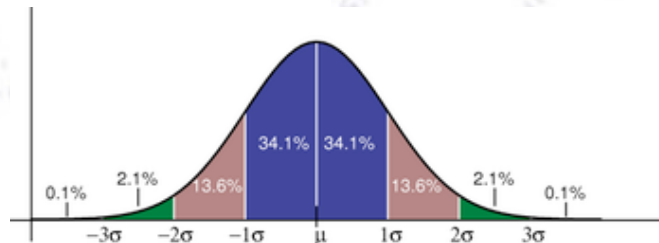# Applied ML

Loss Functions &
Stochastic Gradient Descent

Troels C. Petersen (NBI)

"Statistics is merely a quantisation of common sense - Machine Learning is a sharpening of it!"
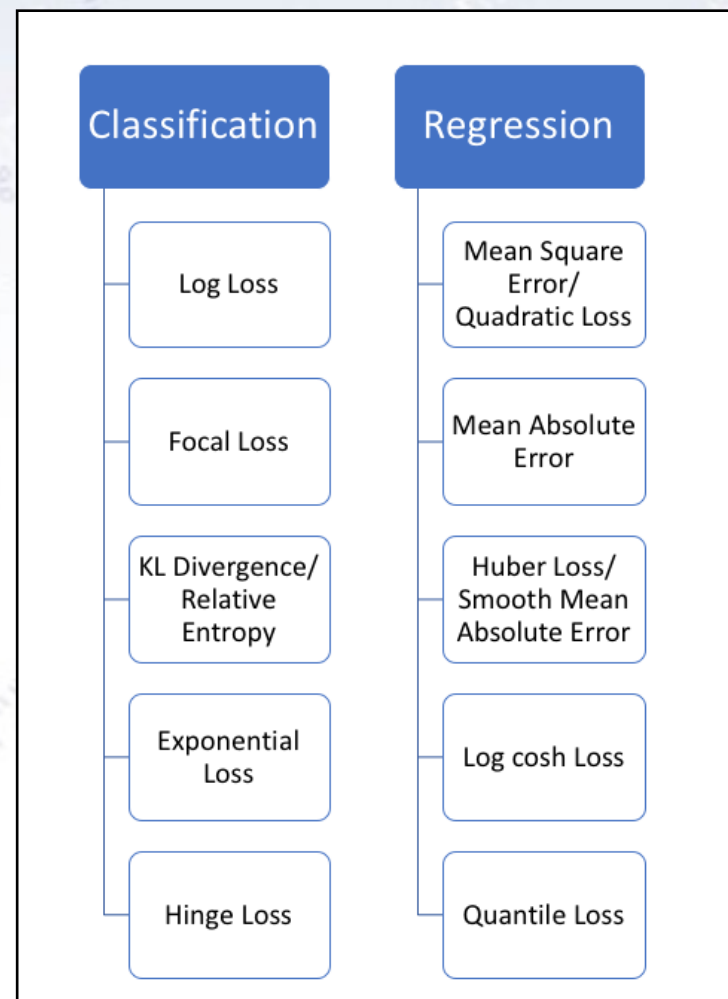
What loss function to use?

The choice of loss function depends on the problem at hand, and in particular what you find important!

In classification:
- Do you care how wrong the wrong are?
- Do you want pure signal or high efficiency?
- Does it matter what type of errors you make?

In regression:
- Do you care about outliers?
- Do you care about size of outliers?
- Is core resolution vital?

Classification	Regression
Log Loss	Mean Square Error/ Quadratic Loss
Focal Loss	Mean Absolute Error
KL Divergence/ Relative Entropy	Huber Loss/ Smooth Mean Absolute Error
Exponential Loss	Log cosh Loss
Hinge Loss	Quantile Loss

What loss function to use?

The choice of loss function depends on the problem at hand, and in particular what you find important!

Loss functions for classification

Classification	Regression
Log Loss	Mean Square Error/ Quadratic Loss
Focal Loss	Mean Absolute Error
KL Divergence/ Relative Entropy	Huber Loss/ Smooth Mean Absolute Error
Exponential Loss	Log cosh Loss
Hinge Loss	Quantile Loss

What loss function to use?

The choice of loss function depends on the problem at hand, and in particular what you find important!

Loss functions for classification

Binary Cross Entropy (aka. LogLoss or Logistic Loss):

$$\mathcal{L} = -\frac{1}{N} \sum_{n=1}^{N} \left[y_n \log \hat{y}_n + (1 - y_n) \log(1 - \hat{y}_n) \right]$$

Classification

Regression

Log Loss

Mean Square Error/ Quadratic Loss

Relative Entropy

Smooth Mean Absolute Error

Exponential Loss

Log cosh Loss

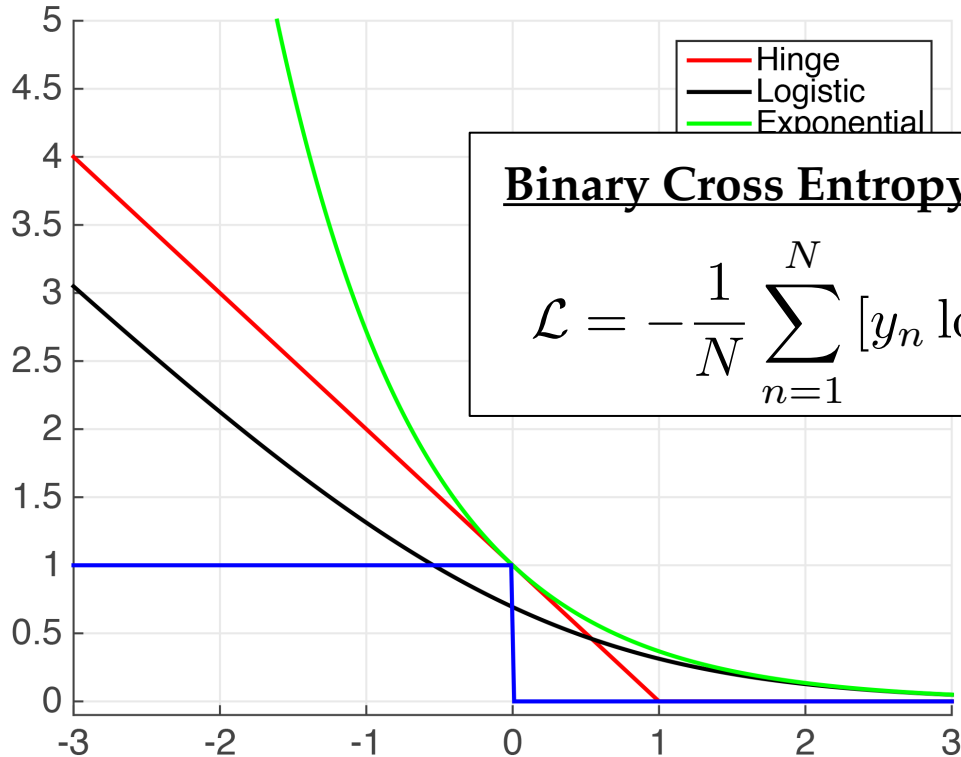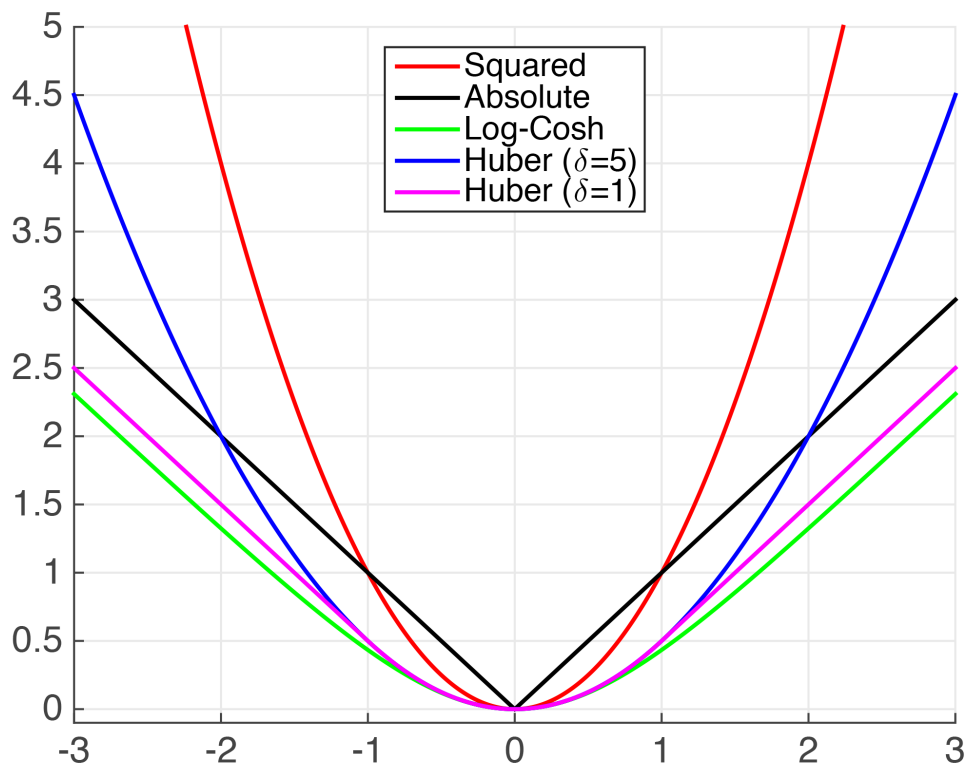Hinge Loss

Quantile Loss

What loss function to use?

The choice of loss function depends on the problem at hand, and in particular what you find important!

Loss functions for regression

Classification	Regression
Log Loss	Mean Square Error/ Quadratic Loss
Focal Loss	Mean Absolute Error
KL Divergence/ Relative Entropy	Huber Loss/ Smooth Mean Absolute Error
Exponential Loss	Log cosh Loss
Hinge Loss	Quantile Loss

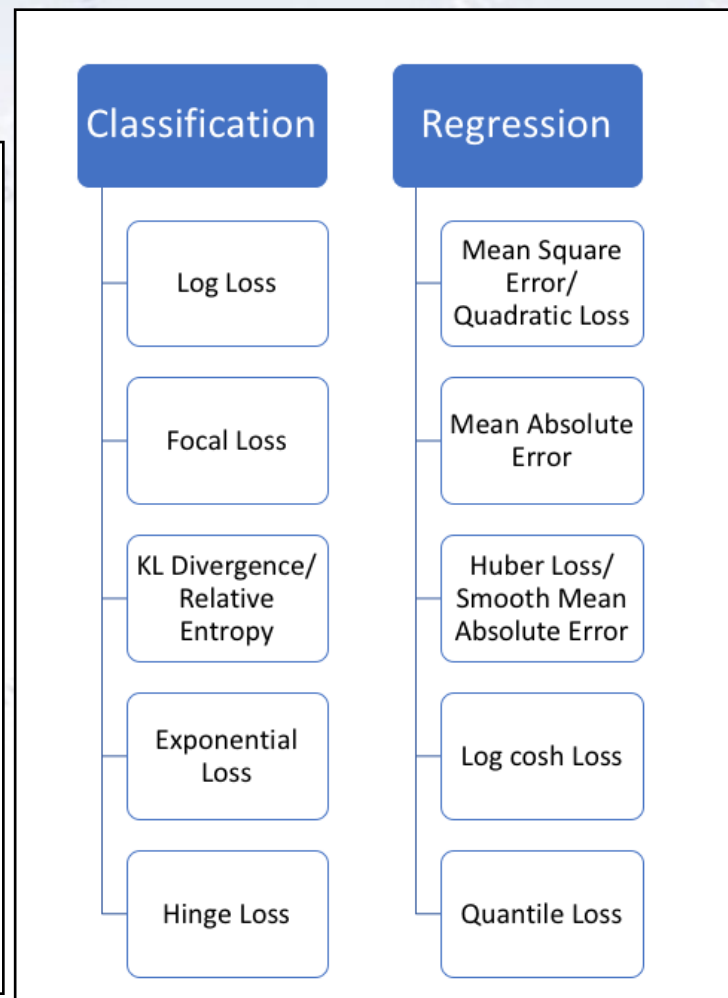Discussion of regression loss functions

What loss function to use?

The choice of loss function depends on the problem at hand, and in particular what you find important!

In classification:
- Do you care how wrong the wrong are?
- Do you want pure signal or high efficiency?
- Does it matter what type of errors you make?

In regression:
- Do you care about outliers?
- Do you care about size of outliers?
- Is core resolution vital?

Ultimately, the loss function should be tailored to match the wishes of the user. This is however not always that simple, as this might be hard to even know!

Classification	Regression
Log Loss	Mean Square Error/ Quadratic Loss
Focal Loss	Mean Absolute Error
KL Divergence/ Relative Entropy	Huber Loss/ Smooth Mean Absolute Error
Exponential Loss	Log cosh Loss
Hinge Loss	Quantile Loss