ROC curves

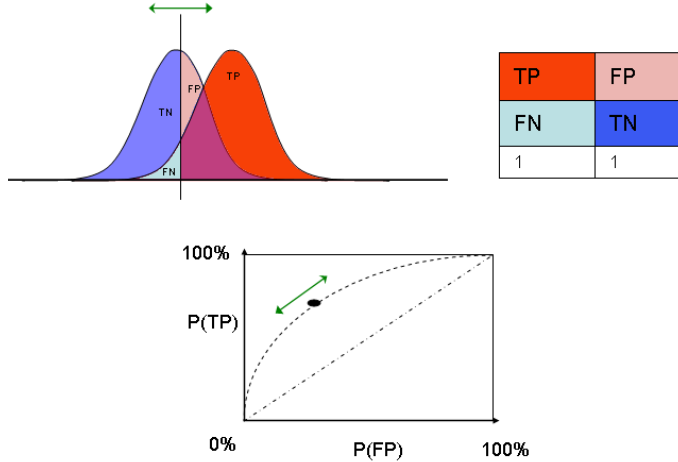**Jeffrey Leek**
Johns Hopkins Bloomberg School of Public Health
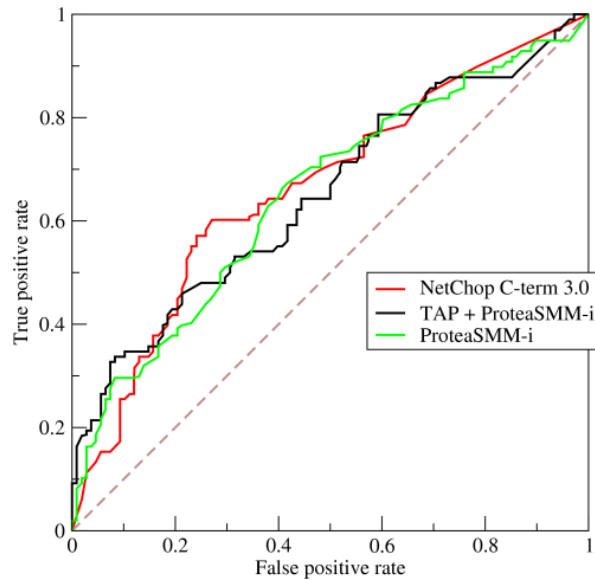
Why a curve?

- In binary classification you are predicting one of two categories

    - Alive/dead

    - Click on ad/don't click

- But your predictions are often quantitative

    - Probability of being alive

    - Prediction on a scale from 1 to 10

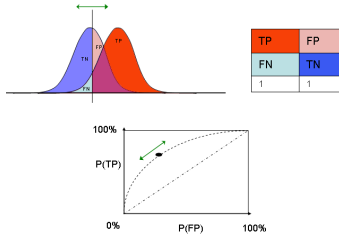- The *cutoff* you choose gives different results

ROC curves

TP	FP
FN	TN
1	1

http://en.wikipedia.org/wiki/Receiver_operating_characteristic

An example

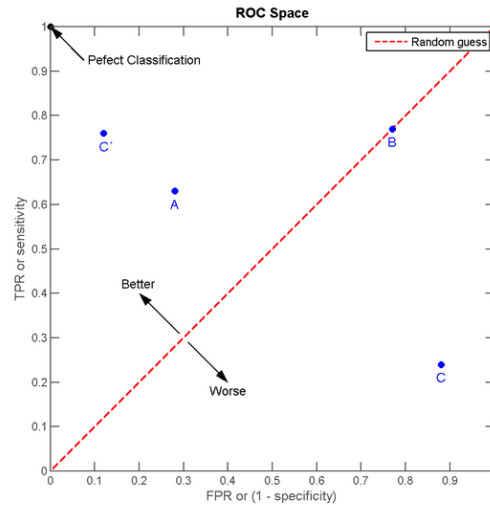http://en.wikipedia.org/wiki/Receiver_operating_characteristic

Area under the curve

- AUC = 0.5: random guessing

- AUC = 1: perfect classifer

- In general AUC of above 0.8 considered "good"

http://en.wikipedia.org/wiki/Receiver_operating_characteristic

What is good?

http://en.wikipedia.org/wiki/Receiver_operating_characteristic