Statistical Analysis of Genetic Data
Math 155, Spring 2008
Jo Hardin
Some helpful R commands

- remember: `library(limma)`
- To normalize your data:
  `backgroundCorrect`
  `normalizeWithinArrays` # note that `normalizeWithinArrays` does the background correction automatically
- You should look at the argument options for the normalization commands. You have the option to do no normalization (`method="none"`).
- boxplots:
  `boxplot(data.frame(norm.data$M))` # boxplot for each array
  `boxplot(data.frame(log2(orig.dat$Gb)))` # boxplot for each array on the green background data (before normalization)
- loess plots:
  `plotMA(norm.data)` # an MA loess plot for an array (or all arrays)
  `scatter.smooth` # a loess plot for any (x,y) arrays
- loess curve/prediction:
  `scatter.smooth(x,y)`
  `l.pred <- loess(y~x)` # note the arguments to loess
  `scatter.smooth(x,y-l.pred)`
- Kolmogorov-Smirnov test of normality:
  `ks.test(norm.data$M[47,],"pnorm",mean(norm.data$M[47,]),sd(norm.data$M[47,]))`
- qq-plot:
  `qqnorm`
  `cor(qqnorm(norm.data$M[47,])$x,qqnorm(norm.data$M[47,])$y)`
- To look at the summary statistics:
  `summary(norm.data$M[47,])`
- To choose random indices:
  `sample(1:20,3)`
  `sample(1:dim(norm.data$M)[1],10)` # note that dim gives the dimensions
  try: `dim(norm.data$M)`
  `dim(norm.data$M)[1]`