Spring 2014 Immunology 206A
Assignment #3: statistics (due Monday, May 5th)

The goal of the assignment is to get a working familiarity of the groups of related statistical tests that can be used to common data questions ("Are my populations different?" "Are my variables correlated?"), and practical exposure to how to implement these tests in R. When working with Big Data, being able to perform a stat many times is a common activity. We'll discuss some of the compensatory measures one needs to take when performing multiple hypothesis testing.

As with the previous assignments, we'll be dividing the different methods out into the class, and you are responsible for teaching your fellow students about the statistical instruments you have been assigned (when it is appropriate, when it is not appropriate), as well as providing working code examples on how to run the test in R. Here's a nice resource from our friends in field biology as a refresher:

http://udel.edu/~mcdonald/statbigchart.html

In addition to the statistical tests, the assignment will include introductory examples of how to simulate hypotheses in R, and some practical examples of how to keep your code organized by building libraries of functions in R. We've built two example functions for you: Dice() and LoadedDice(). You'll find the code

http://www.stanford.edu/class/immunol206a/assignment3-simulator.R
http://www.stanford.edu/class/immunol206a/assignment3-library.R

Again, the code is expected to follow the general format for previous assignments: include your name, comments, command-line active on an input file of data to produce output statistical summaries. Take a look through the previous assignments and adopt techniques that other classmates have applied to make their code as usable as possible.

Assignments for each course member are shown below. You have each been assigned one test per category. For most tests, two or more people have been assigned your test as well. You are welcome to collaborate with them to save time.

	Descriptive	Test of Effect	Goodness of Fit	Independence	Correlation
Marta	median	one-way anova, model II	g-test	g-test of independence	multiple regression
Erika	confidence interval	nested anova	exact test	chi-square test of indepdence	analysis of covariance
Grace	standard deviation	two-way anova	randomization	g-test of independence	Linear regression
Marvin	mean	unpaired t-test	chi-squared	fisher's exact test	polynomial regression
Zina	variance	one-way anova, model I	exact test	randomization test of indepdence	sign test
Miriam	range	paired t-test	Chi-squared	chi-square test of indepdence	Spearman rank
Winn	standard error	nested anova	randomization	cochran-mantel-haenzel test	Pearson
Cesar	median	two-way anova	g-test	randomization test of indepdence	Spearman rank
Eden	confidence interval	unpaired t-test	randomization	fisher's exact test	Linear regression
Jonathan	standard deviation	paired t-test	exact test	g-test of independence	polynomial regression
Theo	mean	receiver-operator curve	Chi-squared	cochran-mantel-haenzel test	analysis of covariance
Winn	variance	Kaplan-Meyer survival curve	g-test	chi-square test of indepdence	Kruskal-Wallis test