**Things you should probably know:**

Simple probability problems – Homework 2 and 3

The difference between a PDF, PMF, and a CDF

Central Limits Theorem

Definition of a Random Variable

Measures of location (mean, median, mode)

Measures of spread (standard deviation, range)

The parameters of the following distributions, whether they are discrete or continuous, and an example of what type of data might fit these distributions:

Binomial

Poisson

Uniform

Normal

What is a standard normal distribution?

Should be able to calculate and explain

Standard Error

95% Confidence Interval

Coefficient of Variation

Standard Deviation and Variance

Mean

Be familiar with the following statistical distributions and how they are related:

t

F

χ2

Basic data transformations, and when and why they should be applied.

Covariance vs. Correlation

What is Fisher’s r to z transformation? When and why is it used?

What is a linear model?

What is ‘dummy’ coding?

How to calculate the coefficient of determination.

How is the slope calculated in a simple linear regression?

How is the intercept calculated in a simple linear regression?

What is meant by “model selection”?

Methods for model selection

sequential vs. marginal sums of squares.

What is collinearity, why is it a problem, how can we test for it?

What is Maximum Likelihood?

What is a likelihood ratio test, and what determines the number of degrees of freedom?

The statistical hypothesis of a paired t-test

The statistical hypothesis of a t-test

Blocked experimental design

Post-hoc tests following an ANOVA

What does a significant interaction term indicate? What should be done if it is significant?

What is meant by ‘power’?

Beta-coefficients

Odds ratios and the logit transformation

Logistic Regression

General linear model vs. generalized linear model

AIC and BIC

Be able to interpret loadings and PC scores

Why p-values should, or should not, be worshipped.