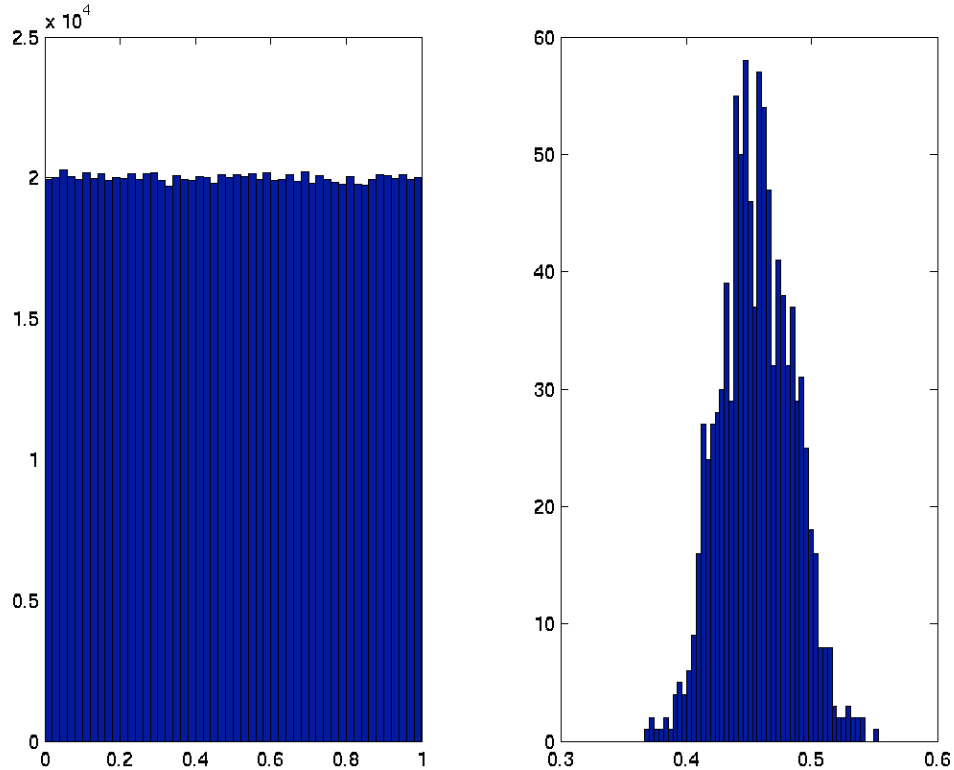
Psych 254
Exercise 4
Due Oct 31

Please email your answers to me, with the title "Psych 258 Exercise 4". The programs should be named "ex4_problemX_<name>.m" where X is the problem # and <name> is your last name.

Problem 1. In this problem you will computationally confirm the central limit theorem. As you may remember from statistics, the central limit theorem states that the distribution of the sample mean will approach the normal distribution with large sample size, regardless of the underlying distribution of the data.

First, create a set of 1,000,000 uniform random variates; these will be our population. Then, cycle through 1000 times, sampling 100 of those variates each time and computing the mean of each sample. Then create a histogram of both the initial distribution of the random variates, and of the sample means across the 1000 runs (use 50 bins in each histogram). You should plot these two graphs in separate panels of a single figure using the subplot command. It should look like this:



Second, use another non-normal distribution, the exponential distribution, to create the random variates. The inverse CDF for the exponential is defined as:

X = -log (1-p)/lambda

where p is the cumulative probability (from zero to one) and lambda is a shape parameter for the distribution. For this exercise, use an exponential distribution with the parameter lambda = 0.5. Again, plot both histograms (the random variates and the sample means) beside each other. Does the sampling distribution look any different than it did for the uniform distribution case?

Problem 2. In this problem you will write a program that demonstrates the shape of the density function (PDF) for various distributions. Your program should do the following:

1. Ask the user which of three distributions they want to plot: standard normal, Student's T, or chi-squared (the user should enter N, T, or X to designate their choice). It should test the input, such that if the subject enters anything other than these three letters it will give a message and ask again for the choice until the subject enters one of these three letters.

2. Ask the user for the appropriate parameters, depending upon which distribution they have chosen. The distributions take the following parameters:

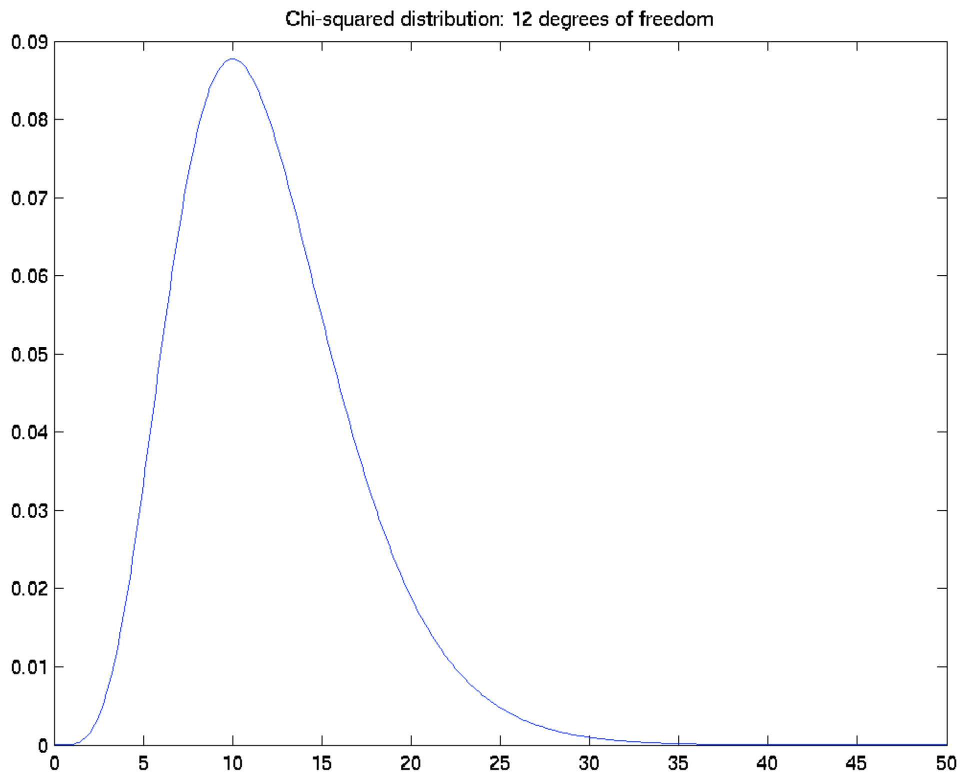
standard Normal – mean=0 and standard deviation=1
Student's T – degrees of freedom
Chi-squared – degrees of freedom

Your program should ask the subject how many degrees of freedom to use if they choose the T or chi-squared distributions, and test this input to make sure that it is reasonable. For the normal distribution you should assume mean=0 and std=1.

The distributions also have different ranges over which they are defined. The normal and T distributions are defined over (-Inf, Inf), but for the purposes of this exercise you can evaluate them over the interval [-6, 6]. The chi-squared distribution is defined over [0, Inf), but for the purposes of this exercise you can evaluate it over the interval [0,50].

3. Use the eval() function to create the appropriate call to the PDF function for the distribution that was chosen. The appropriate functions are normpdf, chi2pdf, and tpdf (for normal, chi-squared, and T distributions respectively); there are equivalent SPM functions if you need to do this on a system that does not have the MATLAB Statistics Toolbox. The call should evaluate the PDF over the appropriate range at 0.1 increments.

4. Plot the PDF, and label it with the name of the distribution and the details (mean and std for normal distribution, df's for T and chi-squared). Use eval() to create the plotting command. The output should look like this:



Chi-squared distribution: 12 degrees of freedom
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
20
25
30
35
40
45
50