Central limit theorem

Artin Armagan

Lab assignment Two

February 8, 2009

Lab outline

The lab will be due on the 20th of February. You will be expected to write a short: 1-2 page lab report and also turn in code. The 1-2 pages do not include plots or graphs that illustrate ideas.

Central limit theorem

Theorem

Let $X_1, ..., X_n$ be a random sample from a distribution with mean μ and variance σ^2. Define

$$\bar{X} = \frac{1}{n} \sum_i x_i.$$

The CLT states that as n goes to infinity that

$$\bar{X} \sim No(\mu, \sigma^2/n).$$

Assignment

We are going to numerically study the CLT for a variety of distributions.
Example code will be provided for you to build on.

The CLT for a binomial distribution

Open matlab
Download the file
drawb.m
run the command
$\gg$drawb(100,.5,100,2000)
you should see a plot of 2000 sums of 100 rv's of binomials and the normal by the CLT for this case
play with p and n

The CLT for a Poisson distribution

Open matlab
Download the file
drawp.m
run the command
$\gg$drawp(100,30,1000)
you should see a plot of 1000 sums of 100 Poisson rv's and the normal by the CLT for this case
play with λ

Breaking the CLT

Open matlab
Download the file
drawcau.m
run the command
$\gg$drawcau(20,200)
you should see a plot of 200 sums of 20 rv's drawn from the distribution

$$p(x) = \frac{1}{x^2}, \quad x > 1,$$

this is like a Cauchy distribution.
This distribution has $\sigma^2 = \infty$, the variance is infinity so the CLT does not hold.
Play around with n (the number of rv's summed).

Specific assignment

Write a simple matlab code to do the following:

1. Consider a fair die (dice) rolling experiment.
   1. Roll a die $10,000$ times and get a histogram of the values observed. What does it look like?
   2. Roll two dice $10,000$ times and sum (or average) the observed values. Now get a histogram for the sums (averages) obtained. What do you observe?
   3. Repeat the above step rolling 3, 10, 50, 100 dice at a time.
   4. For each of the multiple die rolls above, obtain the normal approximation due to the CLT and superimpose the approximated normal pdf over the histograms obtained.

Specific assignment cont'd

2 Now consider a fair coin tossing experiment.

1. Toss a coin $10,000$ times and get the histogram of observed values (H=1, T=0). What does it look like?
2. Now toss two coins $10,000$ times and sum (or average) the observed values. Now get a histogram for the sums (averages) obtained. What do you observe?
3. Repeat the above step tossing 3, 10, 50, 100 coins at a time.
4. For each of the multiple coin tosses above, obtain the normal approximation due to the CLT and superimpose the approximated normal pdf over the histograms obtained.
5. This experiment would be an empirical confirmation of what approximation you have learned?