STA2001 Probability and Statistics I
Computer-based Exercise 9

Pooi Yuen KAM

Problem 1. The Random-Walk and The Central Limit Theorem

The random-walk is a very good model for many physical processes such as the fluctuations of the stock market prices, noise in electronic components, phase noise in lasers, and so on. In the literature, it is also known as a Brownian motion. Here, we will demonstrate that it can easily be generated using software on a computer, and that it asymptotically obeys the Central Limit Theorem.

Imagine a particle that at each time point moves with equal probability, either to the left by a fixed step of size $+1/N$, or to the right by a fixed step of size $-1/N$. After N time points, it ends up at a random point X somewhere within the interval $[-1, +1]$. We can experimentally determine on the computer the probability density function of the random variable X as follows.

Each step to the left or the right can be thought of as the result of the toss of a fair coin. Thus, we can express X as

$$X = \frac{1}{N} \sum_{i=1}^{N} S_i$$

Here, $\{S_i\}_{i=1}^{N}$ is a sequence of Bernoulli random variables, i.e., a sequence of mutually independent, identically-distributed, binary random variables, with $P(S_i = +1) = P(S_i = -1) = 0.5$, for all $i = 1, \cdots, N$. You can generate the sequence $\{S_i\}_{i=1}^{N}$ from the random number generator of your computer, and each sample sequence produces one sample run of the random-walk. Assume that M sample runs of the random-walk are generated.

Divide the interval $[-1, +1]$ into a large number $K+1$ of sub-intervals that you choose. Taking K to be even, we have the sub-intervals $k = -K/2, \cdots, -1, 0, +1, \cdots, +K/2$. For each sub-interval k, let m_k be the number of times that X falls into that sub-interval out of the M sample runs of the random-walk. Plot a histogram of the relative frequency $p(k) = m_k/M$ of X hitting the k-th bin (note that this partition of interval can be done automatically in Python). Choose $K = 100$, and N and M to be at least 1000. On the histogram that you generated, plot the probability density function (pdf) of a Gaussian random variable with the same mean and variance as X. By keeping M fixed and increasing N, you can see the Central limit Theorem coming into play. Repeat the experiment for various values of N and M.

Take note of the definition of the pdf when plotting it on a histogram.