Probability and Statistics Spring 2023
HW7 Matlab assignment

1. The Irwin-Hall distribution

In this problem, you are guided to perform simulations of the Irwin-Hall distribution. The Irwin-Hall distribution is also known as the uniform sum distribution. By its name, the Irwin-Hall distribution is the distribution generated by sum of identical uniform distributions. In addition, the Irwin-Hall distribution is a good example to demonstrate the central limit theorem, which will be covered in Chapter 8.

Suppose that $U_1, U_2, U_3 \cdots U_n$ are independent random variables, each with a continuous uniform distribution on the interval [0, 1] (i.e., the standard uniform distribution). Then, $X_n = \sum_{i=1}^{n} U_i$ follows the Irwin-Hall distribution of order n, where n is a positive integer. The mean and variance of X_n are $\frac{n}{2}$ and $\frac{n}{12}$, respectively. Obviously, for n=1, X_n follows a uniform distribution. For n=2, X_n follows a triangular distribution. When n is large, X_n can be approximated as a normal distribution with mean and variance of $\frac{n}{2}$ and $\frac{n}{12}$.

1.(a) To perform a numerical simulation, please write a Matlab function that generates samples from an Irwin-Hall distribution. Please start with the standard uniform distribution and generate 10^6 random sample points (i.e., Output is a vector with 10^6 elements.). Then, repeat the process for n times and sum up those n vectors. The summation results in a single vector with 10^6 elements, which is a sample of an Irwin-Hall distribution of order n with sample size of 10^6. Your function should have an input parameter n and output of a vector with 10^6 elements.

[Hint: The rand function is convenient to generate samples for the standard uniform distribution.]

1.(b) Now, please use the Matlab function in 1(a) to simulate X_n when n=1, 2, and 20 (i.e., simulate these three cases). Please plot relative frequency histogram of each output vector, which representing samples of X_n, when n=1, 2, 20 (in three separate plots). Then, for each plot, plot a normal distribution with mean and standard deviation of the corresponding X_n. Overlay relative frequency histogram from a sample of X_n with a normal distribution. Please comment on the errors when using a Irwin-Hall distribution to approximate a normal distribution with different values of n.

[Hint: Pease use histogram(Sample_of_Xn,100,'Normalization','pdf') so that histograms are normalized to values of probability density function for ease of comparison with a normal distribution.]

Reference:

https://www.randomservices.org/random/special/IrwinHall.html