Homework 2: Due Thursday February 14, 2019

1. Solve exercises 8.2, 9.2, 12.1 (without the absolute value, i.e. estimate $P(\max_{1\le k\le n} S_k > 2\sqrt{n})$), and 12.3 from the textbook.

2. Write a code to generate $\{S_k : 1 \le k \le n\}$ from exercises 12.1 and 12.3 and use $n = 10,000$ to check that your theoretical estimates are indeed good. Namely, for example for 12.1, generate say 10,000 samples of $\{S_1, \dots, S_{10,000}\}$ and compute the fraction of samples for which $\max_{1\le k\le 10,000} S_k$ was larger than $2\sqrt{10,000} = 200$.

3. Write a code that does the following:

a) Given an integer $N \ge 1$ (a parameter in your code), generate and plot a sample of an approximation to Wiener's Brownian motion:

$$W_N(t) = Z_0 t + \frac{\sqrt{2}}{\pi}\sum_{j=1}^{N}\frac{Z_j}{j}\sin(\pi j t), \text{ for } 0 \le t \le 1,$$

where $\{Z_j : 0 \le j \le N\}$ are i.i.d. standard normals. (We know this converges (almost surely!) to Brownian motion as $N \to \infty$. To demonstrate the code, use $N = 1000$.)

b) For a given $T > 0$ (another parameter in the code), generate and plot a sample of an approximation of Brownian motion

$$B_N(t) = \sqrt{T}W_N(t/T), \text{ for } 0 \le t \le T.$$

(To demonstrate the code, use $T = 2$ and $T = 10$.)

4. Write a code that does the following: For a given $T > 0$ and a given integer $n \ge 1$ (two parameters in the code), generate $\{X_k : 1 \le k \le \lfloor nT \rfloor\}$ i.i.d. standard normal random variables. (Here, $\lfloor x \rfloor$ is the round-down of x.) Then generate and plot a sample of an approximation of Brownian motion:

$$B_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt \rfloor} X_k, \text{ for } 0 \le t \le T.$$

Use $n = 10,000$ and $T = 1$, $T = 2$, $T = 10$, to demonstrate your code.

5. How should the formula for B_n be changed if we did not use standard normals? E.g. what does it become if we use instead a sequence of i.i.d. uniform(-1,1)? What about a sequence of i.i.d. random variables that each take the values ± 1 equally likely? What about if the random variables are not centered? E.g. if we use a sequence of i.i.d. uniform(0,1) random variables?

6. Check the arcsine law with a simulation. That is, generate say 10,000 samples of $\{W_N(t) : 0 \le t \le 1\}$ (say with $N = 1000$) and for each sample find the point $\tau \in [0, 1]$ at which W_N reached its maximal value (for the first time, if there are ties). Then plot the cumulative histogram of the 10,000 samples of τ. This should approximate the CDF of τ. Now, on the same graph (with a different color) plot the actual CDF (coming from the arcsine law). Are they close?

7. Generate four data sets of 100 i.i.d. samples each from the following distributions: exponential(2), normal(6,2) (mean 6, variance 2), Cauchy, and gamma(5,1) (scaling is 1 and shape exponent is 5).

a) For all but Cauchy, plot the pdf from a parameter estimation. Use exponential(θ), normal(μ,σ^2), and Gamma(θ,1).

b) Plot a histogram for each of the data sets, varying the starting point and the smoothing parameter to study their effect.

c) Plot a kernel estimator for each of the data sets, using the kernels: $0.5e^{-|x|}$ and $e^{-x^2/2}/\sqrt{2\pi}$ and the two values 0.5 and 0.1 for the smoothing parameter. (16 plots total)