Exercise 1 (theory) (3 points)

Let $X$ be a random variable with zero mean and variance $\sigma^2$. Using Markov’s inequality, show that

$$P(X \geq t) \leq \frac{\sigma^2}{\sigma^2 + t^2} \quad \forall t \geq 0.$$  

Exercise 2 (theory) (1 + 3 points)

Let $X \sim \text{Exp}(\lambda)$ with $\lambda > 0$.

a) Provide a bound for $P(|X - \mathbb{E}X| \geq \varepsilon)$ as a function of $\lambda$ and $\varepsilon$.

b) Calculate the above probability as a function of $\lambda$ and $\varepsilon$. How big is the difference between the bound and the exact probability for $\lambda = 3$ and $\varepsilon = \frac{1}{2}$?

Exercise 3 (theory) (4 + 3 + 3 points)

Consider a sequence of i.i.d. random variables $Z_1, Z_2, \ldots$ with $P(Z_k = 1) = p$ and $P(Z_k = -1) = 1 - p$ and define

$$X_n = \sum_{k=1}^{n} Z_k \quad \forall n \in \mathbb{N}.$$  

Then $\{X_n\}_{n \geq 0}$ is called a random walk with parameter $p$ starting at 0. Now suppose you want to estimate the probability $\ell = P(X_n = n)$ using the standard estimator

$$\hat{\ell} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}\{X_n^{(i)} = n\},$$  

where $\{X_n^{(1)}\}_{n \geq 0}, \{X_n^{(2)}\}_{n \geq 0}, \ldots$ are i.i.d. copies of $\{X_n\}_{n \geq 0}$.

a) For $n = 5$, use Chebyshev’s inequality to find the minimum sample size so that the error is less than or equal to $10^{-3}$ with probability 0.99.
b) Say you want to ensure that the error of your estimator is, with probability 0.99, smaller than \(0.01 \times \ell\). Using Chebyshev’s inequality, find \(N_{\text{min}}(n)\), the minimum sample size for an arbitrary \(n\). What is the rate of growth of this sample size as \(n \to \infty\) (i.e., what is the order of \(N_{\text{min}}(n)\))?

\[c)\] If, instead, we chose the sample size \(N_{\text{min}}(n)\) using the central limit theorem, what would it be?

Exercise 4 \hspace{1em} (programming) \hspace{1em} (3 + 3 + 4 points)

Solve the following programming exercises without using built-in functions for random number generation (not even \texttt{rand}).

\[a)\] Write a Matlab function \texttt{myrand} generating your own standard uniformly distributed pseudo-random numbers using the Wichman-Hill generator with parameters \(a_1 = 171, a_2 = 170, a_3 = 172\) and \(m_1 = 30269, m_2 = 30307, m_3 = 30323\). Your program should have two parameters \(m\) and \(n\) to specify the size of the matrix which is returned (like \texttt{rand}) and choose seeds based on the current time (for example, using \texttt{clock}).

\[b)\] Write a Matlab function \texttt{mybetarnd} to sample from the beta distribution with parameters \(p, q > 1, p + q > 2\), which has the density

\[f(x; p, q) = \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1} I\{x \geq 0\}.\]

Here \(B\) denotes the beta function (\texttt{beta} in Matlab). Your function should use the acceptance-rejection method based on \(a)\) and have 4 parameters, two for the size \((m\) and \(n\)) and two for the parameters \(p\) and \(q\). Make sure you use the optimal scaling constant \(C\).

\[c)\] Santa wants to bring gifts to the kids in a small village. However, Rudolph ate the list with the naughty and nice kids in this village, so Santa decided to just throw a random number of gifts into each chimney. If there are \(n\) kids in a house, the number of gifts he gives to this house is binomially distributed with parameters \(n\) and \(P\). The gift probability \(P\) for each house is random (depending on Santa’s quickly changing mood) and is independent of the probabilities at other houses. Each \(P\) follows a beta distribution with parameters \(p = 7\) and \(q = 2\). In our particular village, there are five houses with 2, 3, 4, 5 and 6 kids, respectively.

Estimate the expected number of kids who stay without gift and the probability that there is at least one house where the kids get no gifts at all. Use a sample size of at least \(N = 10^4\) and estimate the standard errors of your estimators.

Merry Christmas and a happy New Year! Enjoy your holidays! :-}