HW02

DUE WEEK 3-4

Reference: Foundations of Data Science by Blum, Hopcroft, and Kannan [BHK]

Reading: Sections 12.5, 12.6, 12.7 in [BHK] and lecture notes. Review (discrete and continuous) random variables, sample space, events, independent and dependent r.v.'s, probabilistic inequalities.

1. Alice and Bob play checkers often. Alice is a better player, so the probability that she wins any given game is 0.7, independent of all other games. They decide to play a tournament of n games. Bound the probability that Alice loses the tournament using a Chernoff bound.

2. We have a standard six-sided die. Let X be the number of times that a 6 occurs over n throws of the die. Let p be the probability of the event $X \geq n/4$. Compare the best upper bound on p that you can obtain using Markov's inequality, Chebyshev's inequality, and Chernoff bounds.

3. Let $X_1, \dots, X_n$ be independent $\{-1, 1\}$-valued random variables. Each X_i takes the value 1 with probability 1/2 and else -1. Let $S = \sum_{i=1}^{n} X_i$.

   (a) Let Y be a random variable defined as $Y = |S|$. Prove that Markov's inequality holds for Y.

   (b) Prove Chebyshev's inequality for the above random variable S.

   (c) Show that for any $a > 0$,

   $$\mathbb{P}(S \geq a) \leq 2e^{-a^2/2n}.$$

   Hint: For $t > 0$, show that $\mathbb{E}[e^{tX_i}] = \cosh(tX_i)$. Then, use the fact that $\cosh(x) \leq e^{x^2/2}$ for $x \in [0, 1]$.